

**Sami Hyrynsalmi
Jürgen Münch
Kari Smolander
Jorge Melegati (Eds.)**

LNBIP 500

Software Business

**14th International Conference, ICSOB 2023
Lahti, Finland, November 27–29, 2023
Proceedings**


 **Springer**


OPEN ACCESS


Lecture Notes in Business Information Processing

500


Series Editors

Wil van der Aalst , *RWTH Aachen University, Aachen, Germany*

Sudha Ram , *University of Arizona, Tucson, AZ, USA*

Michael Rosemann , *Queensland University of Technology, Brisbane, QLD, Australia*

Clemens Szyperski, *Microsoft Research, Redmond, WA, USA*

Giancarlo Guizzardi , *University of Twente, Enschede, The Netherlands*

LNBIP reports state-of-the-art results in areas related to business information systems and industrial application software development – timely, at a high level, and in both printed and electronic form.

The type of material published includes

- Proceedings (published in time for the respective event)
- Postproceedings (consisting of thoroughly revised and/or extended final papers)
- Other edited monographs (such as, for example, project reports or invited volumes)
- Tutorials (coherently integrated collections of lectures given at advanced courses, seminars, schools, etc.)
- Award-winning or exceptional theses

LNBIP is abstracted/indexed in DBLP, EI and Scopus. LNBIP volumes are also submitted for the inclusion in ISI Proceedings.


Sami Hyrynsalmi · Jürgen Münch ·
Kari Smolander · Jorge Melegati
Editors


Software Business

14th International Conference, ICSOB 2023
Lahti, Finland, November 27–29, 2023
Proceedings

Editors

Sami Hyrynsalmi 
LUT University
Lahti, Finland

Kari Smolander 
LUT University
Lappeenranta, Finland

Jürgen Münch 
Reutlingen University
Reutlingen, Germany

Jorge Melegati 
Free University of Bozen-Bolzano
Bolzano, Italy



ISSN 1865-1348

ISSN 1865-1356 (electronic)

Lecture Notes in Business Information Processing

ISBN 978-3-031-53226-9

ISBN 978-3-031-53227-6 (eBook)

<https://doi.org/10.1007/978-3-031-53227-6>

© The Editor(s) (if applicable) and The Author(s) 2024. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

Welcome to the proceedings of the 14th International Conference on Software Business (ICSOB 2023). This edition of the conference was hosted in the vibrant city of Lahti, Finland, from November 27 to 29, 2023.

This edition of the conference was hosted by Lappeenranta-Lahti University of Technology (LUT University). Established in 1969, LUT University is a prominent Finnish public research institution with a rich history of academic excellence. The university's Lappeenranta campus graces the picturesque shores of Lake Saimaa, Europe's fourth-largest lake, while its second campus is nestled in the vibrant city of Lahti. As a University of Technology, LUT University specializes in engineering and technology. With a dedicated team of 1,237 staff members and a student body of 7,110, the university cultivates a vibrant academic community.

The conference brought together researchers and practitioners in the field to explore the theme "Digital Agility: Mastering Change in Software Business and Digital Services" and addressed the challenges of managing and leading software-intensive businesses in the relentless pace of technological change and the paramount need for innovation.

The response to this year's conference was record-breaking, with a total of 100 submissions across various categories. We were delighted to announce that out of 79 research track submissions, 27 papers were accepted as full research papers, while 8 were accepted as short research papers. The rigorous review process, led by at least three experts for each submission, ensured the high quality and relevance of the papers presented at ICSOB 2023.

In addition to the main conference tracks, we received 8 applications for the PhD retreat accompanying the conference, which provided an invaluable platform for emerging scholars to engage with established researchers and receive valuable feedback on their work. Furthermore, the poster and demo track received 11 submissions, showcasing innovative applications and practical aspects of software business research. We also had two proposals for workshops and tutorials, contributing to the diverse range of activities and discussions at the conference.

The various topics covered at ICSOB 2023 were vast and vital to the evolving landscape of software business. These included "Software Product Management and Development", "Digital Services, Systems, and Transformation", "Software Ecosystems and Platforms", "Software Business Development", and "Startups and New Venture Creation".

As with previous ICSOB conferences, all accepted papers were published in the conference proceedings by Springer in the Lecture Notes in Business Information Processing (LNBIP) series, and we were proud to announce that the proceedings were published with an Open Access (OA) license, ensuring the widest possible dissemination of the valuable insights and knowledge shared during the event.

The conference featured two captivating keynote presentations that enriched our understanding of strategy and innovation in the software business domain. We were

honored to have Paavo Ritala, a distinguished figure in the field, as one of our keynote speakers. Professor Ritala holds the title of Professor of Strategy and Innovation at LUT Business School (LBS). His research encompasses a wide array of critical themes, including ecosystems and platforms, the pivotal role of data and digital technologies in organizations, collaborative innovation, sustainable business models, and the circular economy. In his distinguished keynote address, Professor Ritala provided a comprehensive and in-depth exploration of the most recent breakthroughs and discoveries emerging from his research portfolio with the keynote titled “The Generative AI Paradox: Strategizing in the New Wave of General-Purpose Technologies”.

The conference’s second keynote was delivered by the accomplished Barbara Hoisl, a renowned authority in the field of strategy, and a seasoned consultant with a specialization in Exponential Strategy. Barbara draws from over 30 years of direct, first-hand experience in the global software and Internet industry. Barbara’s keynote presentation, “The Gift of Thinking Big—What Software People Can Give to the World”, shed light on the crucial intersection of strategy, innovation, and the software business domain. These inspiring keynotes greatly enriched our conference experience and expanded our horizons in this dynamic field.

We were happy to see vibrant discussions, collaborations, and discoveries that ICSOB 2023 inspired. On behalf of the organization team, we would like to express our sincere gratitude to the members of the Program Committee and the additional reviewers for their tireless efforts in evaluating the submissions and ensuring the high quality of the conference. The contributions of the Steering and Organizing Committees and all the chairs were of enormous value in building a successful conference. We also extend our gratitude to all the authors who submitted contributions to the conference, all the authors who presented papers, the keynote speakers, the various audiences who participated in very inspirational discussions during the conference, and the practitioners who shared their experiences and thoughts.

We are delighted to have had the opportunity to enhance the visibility of exceptional papers presented at the conference. In recognition of their outstanding quality and significance, the authors of these selected papers were extended a journal invitation. They were encouraged to submit an expanded version of their originally accepted ICSOB paper for inclusion in a Special Issue dedicated to Software Production within the *Information and Software Technology* journal (IST). We firmly believe that these extended papers will deliver substantial and influential contributions to the special issue, further advancing the discourse and knowledge in the field.

This year, we were particularly excited to highlight two pivotal workshops: The first, “Using Hypothesis Engineering to Manage the Software Architecture Evolution in an Environment with Uncertain Requirements”, by Eduardo Guerra and João Daniel, provided an in-depth exploration into the innovative strategies for navigating the complexities of software architecture in the face of evolving and uncertain requirements. This workshop brought together a diverse group of experts to discuss the integration of hypothesis engineering as a pivotal tool for adaptive and resilient software development. The second workshop, “The Value of Digital Twins for Design Thinking in Digital Agility: The Scene2Model Approach”, by Wilfrid Utz and Iulia Vaidian, offered a unique perspective on leveraging digital twin technology to enhance

design thinking in agile environments. It highlighted the transformative potential of the Scene2Model approach, illustrating how digital twins can serve as critical assets in advancing digital agility. We are proud to present the collective knowledge and innovative ideas shared in these workshops, hoping they will inspire and catalyze further progress in our community.

Thank you for being a part of this remarkable journey, and we appreciate the fruitful interactions during ICSOB 2023.

November 2023

Sami Hyrynsalmi
Jürgen Münch
Kari Smolander
Jorge Melegati

Organization

General Chair

Kari Smolander LUT University, Finland

Program Committee Chairs

Sami Hyrynsalmi LUT University, Finland
Jürgen Münch Reutlingen University, Germany

Proceedings Chair

Jorge Melegati Free University of Bozen-Bolzano, Italy

PhD Retreat Chairs

Rodrigo Santos UNIRIO, Brazil
Xiaofeng Wang Free University of Bozen-Bolzano, Italy
Abayomi Baiyere Queen's University, Canada

Poster Chairs

Ella Peltonen University of Oulu, Finland
Kai-Kristian Kemell University of Helsinki, Finland

Companion Proceedings Chair

Andrey Saltan LUT University, Finland

Publicity Chair

Dimitri Petrik University of Stuttgart, Germany

Website Chair

Erno Vanhala LUT University, Finland

Local Organizing Chair

Sonja Hyrynsalmi LUT University, Finland

Student Volunteer Chair

Marianne Seppänen LUT University, Finland

Steering Committee

Krzysztof Wnuk (Chair)	Blekinge Institute of Technology, Sweden
Xiaofeng Wang	Free University of Bozen-Bolzano, Italy
Anh Nguyen Duc	University of South Eastern Norway, Norway
Eriks Klotins	Blekinge Institute of Technology, Sweden
Helena Holmström Olsson	Malmö University, Sweden
Pasi Tyrväinen	University of Jyväskylä, Finland
Georg Herzwurm	Universität Stuttgart, Germany
Jan Bosch	Chalmers University of Technology, Sweden
Slinger Jansen	Utrecht University, the Netherlands
Pekka Abrahamsson	Tampere University, Finland
Noel Carroll	University of Galway, Ireland
Antonio Martini	University of Oslo, Norway
Sami Hyrynsalmi	LUT University, Finland
Mari Suoranta	University of Jyväskylä, Finland
Anna-Lena Lamprecht	University of Potsdam, Germany
João M. Fernandes	University of Minho, Portugal
Ricardo J. Machado	University of Minho, Portugal
Casper Lassenius	Aalto University, Finland
Kari Smolander	LUT University, Finland
Tiziana Margaria	LERO, Ireland
Michael A. Cusumano	Massachusetts Institute of Technology, USA

Bala Iyer	Babson College, USA
N. Venkatraman	Boston University, USA
Björn Regnell	Lund University, Sweden
Inge Van De Weerd	Utrecht University, the Netherlands
Olga De Troyer	Vrije Universiteit Brussel, Belgium

Program Committee

Matthew Ajimati	National University of Ireland Galway, Ireland
Azeem Akbar	LUT University, Finland
Abdullah Aldaej	Imam Abdulrahman Bin Faisal University, Saudi Arabia
Abayomi Baiyere	Queen's University, Canada
Woubshet Behutiye	University of Oulu, Finland
Richard Berntsson Svensson	Chalmers – University of Gothenburg, Sweden
Jan Bosch	Chalmers University of Technology, Sweden
Regina M.M. Braga	Universidade Federal de Juiz de Fora, Brazil
David Callele	University of Saskatchewan, Canada
Noel Carroll	National University of Ireland Galway, Ireland
Joelma Choma	UFSCar Sorocaba, Brazil
Stanislav Chren	Aalto University, Finland
José Maria David	Federal University of Juiz de Fora, Brazil
Andreas Drechsler	Victoria University of Wellington, New Zealand
Henry Edison	Blekinge Institute of Technology, Sweden
Fabian Fagerholm	Aalto University, Finland
João M. Fernandes	University of Minho, Portugal
Emma Forsgren	University of Leeds, UK
Samuel A. Fricker	FHNW, Switzerland
Juan Garbajosa	Universidad Politécnica de Madrid, Spain
Reetta Ghezzi	University of Jyväskylä, Finland
Maren Gierlich-Joas	Copenhagen Business School, Denmark
Javier Gonzalez-Huerta	Blekinge Institute of Technology, Sweden
Emma Gritt	University of Leeds, UK
Paul Grünbacher	Johannes Kepler University Linz, Austria
Eduardo Guerra	Free University of Bolzen-Bolzano, Italy
Georg Herzwurm	University of Stuttgart, Germany
Helena Holmström Olsson	University of Malmo, Sweden
Jukka Huhtamäki	Tampere University, Finland
Sonja Hyrynsalmi	LUT University, Finland
A.K.M. Najmul Islam	LUT University, Finland
Slinger Jansen	Utrecht University, the Netherlands

Antero Järvi	University of Turku, Finland
Jussi Kasurinen	LUT University, Finland
Kai-Kristian Kemell	University of Helsinki, Finland
Petri Kettunen	University of Helsinki, Finland
Dron Khanna	Free University of Bozen-Bolzano, Italy
Kai Kimppa	University of Turku, Finland
Jens Knodel	Caruso GmbH, Germany
Antti Knutas	LUT University, Finland
Jani Koskinen	University of Turku, Finland
Kari Koskinen	Aalto University, Finland
Samuli Laato	University of Turku, Finland
Casper Lassenius	Aalto University, Finland
Ulrike Lechner	Universität der Bundeswehr München, Germany
Hongxiu Li	Turku School of Economics, Finland
Johan Linåker	RISE Research Institutes of Sweden, Sweden
Francesca Lonetti	CNR-ISTI, Italy
Andrey Maglyas	LUT University, Finland
Paulo Maia	State University of Ceará, Brazil
Tiziana Margaria	Lero, Ireland
Gerardo Matturro	Universidad ORT Uruguay, Uruguay
Jorge Melegati	Free University of Bozen-Bolzano, Italy
Tommi Mikkonen	University of Jyväskylä, Finland
Rahul Mohanani	University of Jyväskylä, Finland
Matti Muhos	University of Oulu, Finland
Jiri Musto	LUT University, Finland
Tuomas Mäkilä	University of Turku, Finland
Timo Mäkinen	Tampere University, Finland
Niko Mäkitalo	University of Jyväskylä, Finland
Anh Nguyen Duc	University College of Southeast Norway, Norway
Emil Numminen	Blekinge Institute of Technology, Sweden
Arto Ojala	University of Vaasa, Finland
Pablo Oliveira Antonino	Fraunhofer, Germany
Shola Oyedeji	LUT University, Finland
Maria Paasivaara	LUT University, Finland
Efi Papatheocharous	RISE Research Institutes of Sweden, Sweden
Samuli Pekkola	University of Jyväskylä, Finland
Ella Peltonen	University of Oulu, Finland
Wolfram Pietsch	Aachen University of Applied Sciences, Germany
Tero Päivärinta	Luleå University of Technology, Sweden
Minna Rantanen	University of Turku, Finland
Rodrigo Rebouças de Almeida	Federal University of Paraíba, Brazil
Matti Rossi	Aalto University, Finland

Andrey Saltan	LUT University/HSE University, Finland
Rodrigo Santos	UNIRIO, Brazil
Marko Seppänen	Tampere University, Finland
Pertti Seppänen	University of Oulu, Finland
Dominik Siemon	LUT University, Finland
Kari Smolander	LUT University, Finland
Barbara Steffen	TU Dortmund, Germany
Gero Strobel	University Duisburg-Essen, Germany
Victor Stroele	Federal University of Juiz de Fora, Brazil
Erkki Sutinen	University of Turku, Finland
Kari Systä	Tampere University, Finland
Nirnaya Tripathi	University of Oulu, Finland
Pasi Tyrväinen	University of Jyväskylä, Finland
Michael Unterkalmsteiner	Blekinge Institute of Technology, Sweden
Ville Vakkuri	University of Jyväskylä, Finland
George Valença	Universidade Federal Rural de Pernambuco (UFRPE), Brazil
Rini Van Solingen	Delft University of Technology, the Netherlands
Erno Vanhala	LUT University, Finland
Davi Viana	Federal University of Maranhão, Brazil
Hannu Vilpponen	University of Jyväskylä, Finland
Xiaofeng Wang	Free University of Bozen-Bolzano, Italy
Karl Werder	University of Cologne, Germany
Krzysztof Wnuk	Blekinge Institute of Technology, Sweden
Katariina Yrjönkoski	Tampere University, Finland
Ehsan Zabardast	Blekinge Institute of Technology, Sweden
Markus Philipp Zimmer	Leuphana University Lüneberg, Germany

Additional Reviewers

Ava Heinonen	Leonardo Banh
Lucas Abreu	Anita Hidayati
Luiz Alexandre Costa	Janne Harjamäki
Maha Sroor	Yagmur Turhan
Harri Keto	Dongmei Gao
Rodrigo Zacarias	Damian Kedziora
Rong Huang	Juliana Outão
Sophia Mannina	Paulo Malcher
Shan Feng	Nurbojatmiko Nurbojatmiko
Alexandra Hettich	Reza Toorajipour
Mario Simaremare	

Contents

Requirements

Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review	3
<i>Dimitri Petrik, Anne Untermann, and Henning Baars</i>	
Are Business Expectations Aligned with the Development Plan Made by the Software Architecture Area? A Case Study on Agile Teams in a Large Company	19
<i>Marcelo Augusto da Silva, Inaldo Capistrano Costa, and Eduardo Martins Guerra</i>	
Investigating Open Innovation Practices to Support Requirements Management in Software Ecosystems	35
<i>Paulo Malcher, Davi Viana, Pablo Oliveira Antonino, and Rodrigo Pereira dos Santos</i>	
Requirements Tool Practices that Drive Business Agility	51
<i>Andreas Birk</i>	

Software Procurement

On Public Procurement of ICT Systems: Stakeholder Views and Emerging Tensions	61
<i>Reetta Ghezzi and Tommi Mikkonen</i>	
Improving Communication and Collaboration in Enterprise Architecture Projects: Three Propositions from Three Public Sector EA Projects	77
<i>Ari Rouvari and Samuli Pekkola</i>	
Navigating ICT In-House Procurement in Finland: Evaluating Legal Frameworks and Practical Challenges	92
<i>Reetta Ghezzi, Minnamaria Korhonen, Hannu Vilpponen, and Tommi Mikkonen</i>	
Artificial Intelligence Procurement Assistant: Enhancing Bid Evaluation	108
<i>Muhammad Waseem, Teerath Das, Teemu Paloniemi, Miika Koivisto, Eeli Räsänen, Manu Setälä, and Tommi Mikkonen</i>	

Platforms, Ecosystems and Data

Who Does What? Evolving Division of Responsibilities in a B2B Platform 117
Jaakko Vuolasto

Understanding User Feedback in Software Ecosystems: A Study
on Challenges and Mitigation Strategies 132
Bachan Ghimire, Ze Shi Li, and Daniela Damian

A Survey on Perceptions of Data Sharing in the Norwegian Public Sector 148
*Leif Z. Knutsen, Bertha Ngereja, Ingebjørg Flaata Bjaaland,
Jo E. Hannay, and Sinan S. Tanilkan*

Investigating the Barriers that Women Face in Software Development
Teams Focusing on the Context of Proprietary Software Ecosystems 164
*Juliana Carvalho Silva do Outão, Luiz Alexandre Martins da Costa,
Rodrigo Pereira dos Santos, and Alexander Serebrenik*

Artificial Intelligence

Business and Ethical Concerns in Domestic Conversational Generative
AI-Empowered Multi-robot Systems 173
*Rebekah Rousi, Hooman Samani, Niko Mäkitalo, Ville Vakkuri,
Simo Linkola, Kai-Kristian Kemell, Paulius Daubaris, Ilenia Fronza,
Tommi Mikkonen, and Pekka Abrahamsson*

Prompt Patterns for Agile Software Project Managers: First Results 190
Kari Sainio, Pekka Abrahamsson, and Tero Ahtee

Startup Creation Beyond Hackathons – A Survey on Startup Development
and Support 205
Maria Angelica Medina Angarita, Martin Kolnes, and Alexander Nolte

Starting Collaborations Between SMEs and Researchers in Software
Engineering 222
Sergio Rico, Felix Dobslaw, and Lena-Maria Öberg

Towards a Business Case for AI Ethics 231
Mamia Agbese, Erika Halme, Rahul Mohanani, and Pekka Abrahamsson

What Is the Cost of AI Ethics? Initial Conceptual Framework and Empirical
Insights 247
Kai-Kristian Kemell and Ville Vakkuri

Software Startups

Benefits, Challenges, and Implications of Open-Source Software for Health-Tech Startups: An Empirical Study	265
<i>Noman Ahmad and Nirnaya Tripathi</i>	
Corporate Startups: A Systematic Literature Review on Governance and Autonomy	283
<i>Konstantin Garidis, Alexander Rossmann, and Alan Murray</i>	
Exploring the Finnish Impact Investing Ecosystem: Perspectives on Challenges from Technology Startups	299
<i>Timo Okker, Rahul Mohanani, Tommi Auvinen, and Pekka Abrahamsson</i>	
Practitioner Views on Analytics for Software Startups: A Preliminary Guide Based on Gray Literature	315
<i>Usman Rafiq, Frédéric Pattyn, and Xiaofeng Wang</i>	

Software Product Management

An Evaluation of the Product Security Maturity Model Through Case Studies at 15 Software Producing Organizations	327
<i>Elena Baninemeh, Harold Toomey, Katsiaryna Labunets, Gerard Wagenaar, and Slinger Jansen</i>	
Strategic Digital Product Management in the Age of AI	344
<i>Helena Holmström Olsson and Jan Bosch</i>	
Experimentation in Early-Stage Video Game Startups: Practices and Challenges	360
<i>Henry Edison, Jorge Melegati, and Elizabeth Bjarnason</i>	

Software and Business Co-Development

DevOps Challenges and Risk Mitigation Strategies by DevOps Professionals Teams	369
<i>Nasreen Azad</i>	
Positive Customer Experience is Enhanced by Effective Agile Practices	386
<i>Riina Piironen, Ilkka Jormanainen, and Markku Tukiainen</i>	
Information-Centric Adoption and Use of Standard Compliant DevSecOps for Operational Technology: From Experience to Design Principles	400
<i>Henry Haverinen, Tero Päivärinta, Jussi Vänskä, and Henry Joutsijoki</i>	

Exploring Emotions in Online Team Meetings: Unpacking Agile Retrospective 416
Dron Khanna and Abdullah Aldaej

Emerging Digital World

Feeling the Elephant: Insiders’ Perspectives on the Metaverse 427
Fabrcio de Oliveira, Xiaofeng Wang, and Luciana Zaina

Carbon Footprint Calculations for a Software Company – Adapting GHG Protocol Scopes 1, 2 and 3 to the Software Industry 442
Antti Sipilä, Laura Partanen, and Jari Porras

Understanding Cost Dynamics of Serverless Computing: An Empirical Study 456
Muhammad Hamza, Muhammad Azeem Akbar, and Rafael Capilla

Quantum Software Ecosystem: Stakeholders, Interactions and Challenges 471
Vlad Stirbu and Tommi Mikkonen

Dynamic Capabilities for Sustainable Digital Transformation Amid Crisis: Insights from Law Firms in Emerging Economy 478
Mikhail O. Adisa, Gbadebo A. Ojikutu, Larry Abdullahi, Shola Oyedeji, and Jari Porras

Research Streams of Barriers to Digital Transformation: Mapping Current State and Future Directions 493
Henning Brink, Fynn-Hendrik Paul, and Sven Packmohr

Author Index 513

Requirements



Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review

Dimitri Petrik^{1,2} , Anne Untermann², and Henning Baars² 

¹ Graduate School of Excellence Advanced Manufacturing Engineering (GSaME), Nobelstr. 12, 70569 Stuttgart, Germany

dimitri.petrik@gsame.uni-stuttgart.de

² University of Stuttgart, Keplerstr. 17, 70174 Stuttgart, Germany

{anne.untermann, henning.baars}@bwi.uni-stuttgart.de

Abstract. Organizations must gain insights into often fragmented and isolated data assets and overcome data silos to profitably leverage data as a strategic resource. Data catalogs are an increasingly popular approach to achieving these objectives. Despite the perceived importance of data catalogs in practice, relatively little research exists on how to design corporate data catalogs. It is also obvious that the existing market solutions have to be customized to the specific organizational needs. This paper presents a list of functional requirements for enterprise data catalogs extracted from a systematic literature review. The requirements can be used to frame and guide more specific research on data catalogs as well as for system selection and customization in practice.

Keywords: Data catalog · metadata · metadata management · requirements

1 Introduction

Recent technological developments in cloud provisioning, analytics technologies, and the Internet of Things foster data collection and analytics which in turn create novel opportunities for organizations to gain a competitive advantage [1]. The automotive industry, for instance, is impacted by analytics-based innovations in manufacturing, product design (i.e., connected and autonomous cars), collaborative services, and – based on that – novel business models [2, 3]. In other industries, too, organizations are increasingly trying to monetize their data together with the own employees’ knowledge and are trying to bundle them to knowledge-intensive services [4]. In doing so, refined data acts as a key strategic resource for organizations that supports identifying optimization opportunities and sustainable efficiency gains in business processes [5]. To leverage these opportunities, organizations require integration and harmonization of data within and beyond the organizational boundaries [6].

Consequently, organizations need an overview of distributed data assets to acquire a sufficient understanding of the data inventory already available to fully exploit the potential of refined data [6]. Typically, the available data is fragmented. It is stored in a multitude of disparate IT systems by numerous departments as well as external actors,

resulting in isolated data silos. Data silos are also a significant hurdle to overcome as suppliers, customers, and the manufacturing organizations themselves are trying to form data ecosystems with big data analytics that lead to even more complex data landscapes. Increasing complexity and, at the same time, decreasing transparency about existing data inventories hamper the discoverability of meaningful datasets and obscure important information about the interrelationships of data, as well as collaboration possibilities of actors, remain hidden. The search processes for relevant data have become long and costly [7]. This, in turn, firstly impedes the provision of knowledge services. Secondly, it prevents relevant initiatives e.g., for self-service analytics and data democratization, in which employees of operational departments are directly involved in value creation and empowered to perform analytics and share data assets without dedicated data experts [8, 9].

To overcome these challenges, organizations require robust data management concepts [10]. Data catalogs are established solutions to tackle those [9]. A data catalog is an enterprise system for metadata management and data curation [11]. It functions as a knowledge and collaboration hub, supports organizations in building sovereign data infrastructures in continuously expanding networks [11], and supports data analysts and other data consumers during the search for data sets, storage locations, intended uses, and other essential information, thus ensuring a better understanding of the existing data landscapes [12].

Multiple commercial (e.g., IBM, AWS or Oracle) and open-source (e.g., Apache Atlas) tools for cataloging are available [11, 14]. It needs to be considered that these are designable and customizable systems that usually cannot be applied off-the-shelf and their tailoring and organizational and technical implementation are non-trivial tasks. Despite the criticality of data catalogs for software-intensive business, issues of their design remain largely under-researched [8]. An initial analysis of the current scientific research literature reveals a lack of design-oriented research and results regarding the subject of enterprise data catalogs. Existing literature reviews indicate that the current research literature has so far mainly concentrated on domain-specific “open data” topic e.g. in the realms of government data, research data, or geospatial data, and is therefore not directly applicable to enterprise scenarios [15]. This state reveals a research gap in the design of enterprise data catalogs, especially in the industrial and inter-organizational data ecosystem contexts. Therefore, we ask: *What are the relevant requirements to design enterprise data catalogs?*

Reflecting on the state of research on data catalogs in the enterprise context, confirms the need for further scientific research on the design and implementation of enterprise data catalogs. For this reason, this paper particularly aims to identify and extract functional requirements for enterprise data catalogs from a systematic analysis of the scientific body of knowledge.

2 Data Catalogs and Metadata Management

Enterprise data catalogs are recognized as enterprise information systems to collect, create and maintain contextual information (i.e., metadata) from heterogeneous source systems [15]. They are context-specific digital data directories in which metadata, i.e.,

data about data, for all existing data objects can be stored centrally and managed securely in order to catalog them in a way that adds value [5]. In an enterprise architecture, data catalogs complement other existing systems for working with data. Functional models often see data catalogs as complementary to data lakes and they are sought to ensure that the data lakes remain manageable and do not become data swamps [10, 16]. They are usually stand-alone software systems (as evidenced by the existing software product landscape [11]) that work hand-in-hand with other data-related subsystems of an enterprise data architecture. For instance, while data quality tools specialize in identifying data problems and fixing them (e.g., through format alignment, standardization, cleansing, and profiling) [17, 18], data catalogs can make the qualified data assets accessible to different roles [11]. In the cross-organizational context of data ecosystems, data catalogs function, for example, complementary to data marketplaces, which provide data brokerage services [10], integrated in interoperable data platforms [11, 19]. To conclude, data catalogs are an integral part of data-driven solutions and thus of software-intensive business, supporting business intelligence and analytics within enterprises or a data ecosystem.

In the existing academic research literature, enterprise data catalogs are associated with data democratization. “Data democratization” implies that non-IT employees are given access to existing data sets and are empowered to use them for data-driven purposes [8]. Accordingly, by providing a conceptual structure as well as various data access functions, data catalogs should facilitate **findability, accessibility, interoperability, and reusability** (FAIR principles) of data assets for the different casual and technical (i.e., analytics experts) users to support the democratization of data. In the literature, this is considered one of the core benefits of their deployment. For this purpose, data catalogs can provide appropriate search mechanisms so that users can discover data sets for their specific use cases [8]. A pertinent design of a data catalog should therefore ensure that the different users can find out which data objects are registered and provide consistent descriptions of the data assets and their locations [8, 20]. Therefore, data catalogs simultaneously function as abstractions of various documentation levels and thereby should facilitate a centralized data access point within and across organizational borders (in a setting with a data catalog that supports a data ecosystem) [11]. Once a user has identified appropriate data sets, they should be made accessible directly through the data catalog. Since data catalog implementation aims to make data from different domains and previous data silos available and usable, ensuring the comprehensive quality of data sets scattered in heterogeneous source systems [21], an **assessment of the quality** of the registered objects plays an eminent role, as this is the only way to generate actual added value for the data consumer. The main component of a data catalog to make data searches possible is the so-called **data inventory**, which models and describes the available data supply [8]. Data might be manually captured by users or automatically collected through interactions with the respective source systems; particularly when pre-built metadata models foster a standardized data capture [8, 22]. Another essential aspect of the data inventory is the detailed documentation of the data sequence (also known as **data lineage**). Data lineage describes the ability to trace data records back to their original source, i.e., data provenance [5, 15, 22, 23]. Because data catalogs are intended to replace manual searches, they should be able to consolidate and **automate**

the corresponding processes which are otherwise often time-consuming and inefficient [8, 23, 24].

Since enterprise data catalogs support metadata management, this section also presents the related work on metadata. Metadata includes information about data sets and can be generated either manually by the data creator or automatically by a system. Metadata can include information about the data creator, record contents and contexts, or timestamps of data creation [25]. In data management, metadata is significant in facilitating access, management and sharing of structured and unstructured data [26]. The National Information Standards Organization (NISO) supports this statement and adds that consistently maintained and structured metadata are used, on the one hand, to help users find appropriate data sets in heterogeneous data structures of information systems and, on the other hand, to capture and subsequently share essential information about these data, thereby promoting data understanding and transparency [27]. Three metadata types can be distinguished [27]:

- Descriptive metadata (1) provides information about the content of data sets and makes it easier for data consumers to identify and understand appropriate data objects for their specific use or research purpose. Exemplary metadata elements are titles, descriptions, or keywords.
- Administrative metadata (2) is a collective term for data related to managing or creating data sets and can be divided into three segments: 1. Technical metadata, such as information about the physical structure of the data set, such as file format, software used, or encoding; 2. Legal metadata, such as information about access rights, copyright restrictions, or intellectual property rights; 3. Data provenance metadata, such as information about the lineage, last modifications, and reasons for the creation of the data set. The information provided thus assists users in interpreting the identified datasets.
- Structural metadata (3) represents the relationship and interaction between the sub-elements of the data set, such as the hierarchy levels or foreign-key-relationships.

Other metadata classifications may also be useful for the discovery of data sets. For example, metadata can be divided into business metadata (i.e., information about the business context and policies), operational metadata (i.e., the information generated automatically during data processing, such as the information about data quality), and technical metadata (i.e., information about the data structure such as the data format or scheme) [28, 29]. This classification can be beneficial because business metadata promotes data understanding by technical or non-technical-savvy staff and enhances interdisciplinary exploration and interpretation of data sets, while operational metadata enables the derivation of insights related to quality development, security, and compliance, and technical metadata is used to document data composition and types [23]. The different existing metadata typologies are often interrelated and, therefore, not always generated and documented separately [29]. Finally, it is helpful to reconstruct the lifecycle of data elements through consistent metadata to enable the search of data objects within complex information systems. Thus, metadata promises to provide real economic value when, for example, it is at least partially automated, and previously collected information is reused to avoid redundant or obsolete metadata and streamline the curating process [30]. When metadata is generated in a way that is readable by both

machines and humans, it promotes interoperability and integration of metadata on the one hand, and allows data sets to be described, discovered, and contextualized [25, 27, 30]. To achieve this, enterprise data catalogs represent the information systems to realize metadata documentation and provisioning [24].

3 Methodology

As a literature review aims to synthesize the existing state of knowledge on a selected phenomenon, we consider it to be a suitable research methodology for extracting functional requirements for enterprise data catalogs as a form of codified design knowledge. We follow established guidelines for a systematic concept-centric literature review on a database level [31]. For the definition of the sample of relevant literature sources, we started with an unsystematic literature search on Google Scholar EBSCOhost and ScienceDirect (with the generic search terms “Data Catalog”) which helped us pinpointing more specific search criteria. From the results we refined the following keywords: ‘data catalog’, ‘metadata catalog’, ‘enterprise’, ‘data repository’, and ‘data register’. The publication period was set to 2006–2023 as data catalogs in their current form represent a relatively new concept. Another relevant selection filter was the accessibility of the publications as well as a focus on conference and journal contributions (academic journals, conference papers, or proceedings): We tried to avoid that incomplete texts, non-accessible papers, or non-peer-reviewed articles. In total, we formulated two search terms that we applied separately across the five databases Web of Science, SpringerLink, ACM Digital Library, IEEEExplore, and AISeL:

1. “data catalog*” OR “metadata catalog*”
2. “data catalog*” AND enterprise

This generated a total of 750 hits with the first search term and 11 with the second. After applying the aforementioned filter criteria, the sample for the first search string was 408 papers, and for the second search term 10 papers. After excluding the duplicates, the sample went down to 391 papers. In the next step, the titles and abstracts were manually analyzed to determine whether they fit the research question and indeed have “data catalogs” as their research subject. Articles dealing with data catalogs in the domains of medicine, politics, astronomy or geography were excluded, as they do not deal with corporate and industrial contexts of use of data catalogs. Nevertheless, a few articles from these research areas were retained if they contained information that could be transferred to the entrepreneurial context. Since the titles and abstracts were often not meaningful, we performed diagonal reading to minimize subjectivity. Here, the introductions, the conclusion of the articles, and the figure and table titles used were examined with respect to the inclusion and exclusion criteria. A total of 45 articles remained. After reading the full texts, a backward search resulted in six additional articles. After the full-text screening, additional papers were removed from the sample that for instance only described projects with happened to include data catalogs. The authors discussed each paper of the initial sample, seeking a consensus within the research team to increase the objectivity of the exclusion. In doing so, the final sample was reduced to 21 relevant articles.

Due to the limited amount of scientific literature on data catalogs in the enterprise context, we broadened our search and explicitly included grey literature, esp. White papers and research reports. After all, white papers and practice reports are considered recognized explanations of practice, which can prepare qualitative expertise and recommendations regarding a specific topic in a consolidated manner. Thus, adhering to standard guidelines for including grey literature in systematic literature reviews [32], we have broadened our sample by including only grey literature with high credibility and high outlet control. Our selection criteria exclude marketing documents from tool providers, focusing solely on reports from reputable research institutes or established management consultancies that are known for leveraging software- and data-driven projects. In addition to assessing the authority of the sources, our inclusion of grey literature was also guided by the perceived objectivity of their statements. In this way, three additional publications could be added. Due to length constraints, the literature sample compiled is detailed in an external appendix, accessible via the following URL: <http://bit.ly/49Jbbp5> (Fig. 1 illustrates the sample creation process).

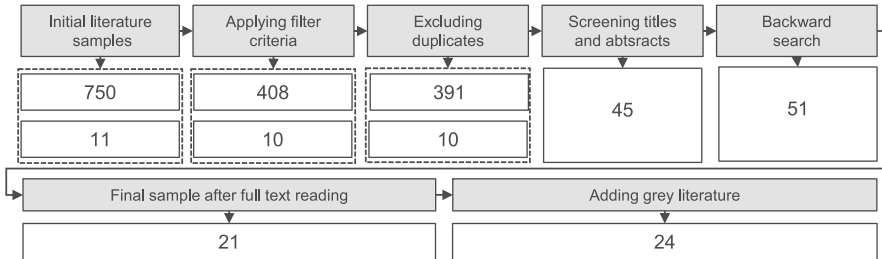


Fig. 1. Illustration of the literature search process and sample creation

During the content analysis of the remaining papers [33], we inductively formed categories for the derivation of functional requirements, guided by the expertise within the research team. According to the inductive technique, the abstraction level is successively increased to develop theory-based main categories from a large number of groupings from the available texts. Each researcher independently reviewed the articles in the created sample, applying coding techniques and labeling the functionalities. These codes were then collectively discussed by the research team to foster a shared understanding and to collaboratively formulate the requirements. In this process, a total of 13 functional requirements were derived.

4 Requirements for Enterprise Data Catalogs

The derived requirements have been grouped into the following six categories, each represented by a unique identifier: metadata management (Requirements R1-4); data inventory (Requirements R5-6), data governance (Requirements R7-9), interoperability (Requirement R10), interface (Requirement R11), collaboration (Requirement R12), intelligent automation (Requirement R13). The requirements were grouped based on

their functional similarity during discussions within the researcher team. Figure 2 integrates the requirements in a functional view on an enterprise data catalog, embedded either in a data lake or in a data platform, based on [11]:

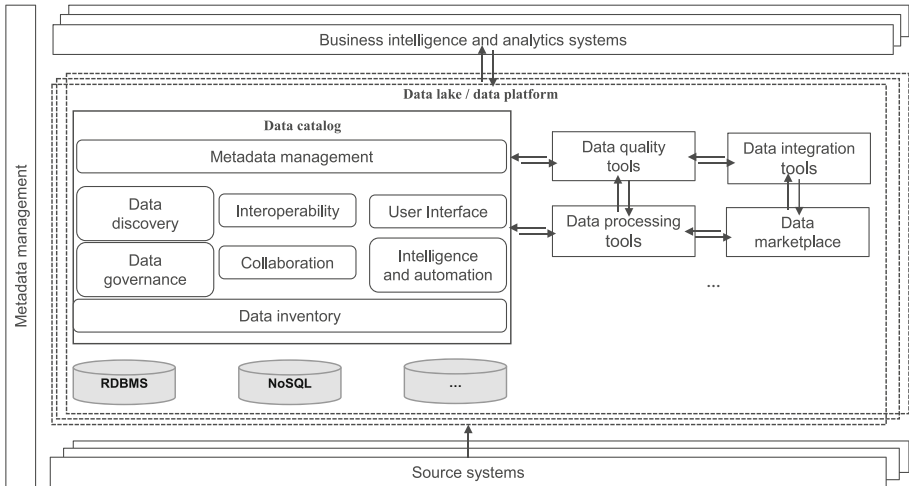


Fig. 2. Functional view on enterprise data catalogs

Data catalogs function as central indexed searchable sources for finding data [8, 24]. To ensure successful and seamless data set searches, robust search functionalities should be integrated into data catalogs that enable users to find data objects for a specific analytics purpose [22, 34]. In particular, the search for keywords, business terms, or metadata should be offered. In addition, using functions that utilize a natural language simplifies the search for data consumers of a non-technical domain [22, 25, 35]. This includes, for example, full-text or semantic search (which is also used in Google searches) to deal with the content of search queries. Designations or titles of data sets, data domains, or business units are first classified and then indexed, resulting in the display of data relating to the content entered [23, 36].

In addition, the role-specific requirements of individual users should be included to avoid missing necessary functionalities or integrating superfluous functions that hinder the search [22]. This results in the following requirement:

R1: *Enterprise data catalogs should be equipped with robust search functionality to enable employees to identify needed data sets by entering, for example, keywords, metadata, or full text, considering role-specific search requirements.*

Furthermore, data catalogs should allow the user to enrich the recorded data objects with complementary information to improve the findability of the data sets and to facilitate the search by giving additional clues about how data objects are related. Finally, high information content promotes the user understanding of data sets and makes data knowledge more consumable. Accordingly, it should ideally be possible to associate data with labels, *identifiers*, and to link them to a searchable source, which provides additional insights into the content and the characteristics of the data [8, 13, 20]. Essential for

indexing is an adequate description of the information about the individual data objects, whereby priority should be given, for example, to the descriptions' completeness, simplicity, and relevance. Based on this information, users can decide whether the data sets are suitable for the respective analytics projects, so the description should be created rather carefully [37]. Tagging functions also improve discoverability significantly [5, 13, 38]: The data is labeled, and it is determined on which level the previously defined metadata variables or attributes are assigned to the respective data. [5, 15]. Depending on the context of the use of the data catalog, data can be tagged at four levels: dataset level (original source dataset), record level (for all data entries in the dataset), entity level (for each data entry), and column level (individual columns in the dataset) [5]. This results in the following requirement:

R2: *Enterprise data catalogs should allow the linking of registered data objects by data providers to adequate identifiers and appropriate indexes to ensure data discovery and facilitate the evaluation of data sets by system users, particularly if the data catalog consolidates data objects from different usage contexts.*

Besides, data catalogs must support metadata documentation while supporting the applicable metadata standards (if applicable). To enable reusability of data objects by aligning enterprise and system-oriented views of data, a complete documentation of metadata should be based on a conceptual (i.e., the context of the creation and the application of the data), a logical (i.e., entities and their relationships to each other as well as associated business objects and attributes) and a physical level (i.e., information related systems, interfaces, data structures and attributes etc.) [8, 21]. Constructive here would be the enrichment of the data with contextual information that can (1) describe the operational context in terms of the domain or subject area in which the data operate, on the one hand, and (2) characterize the technical context through technical details regarding the data source or data set, on the other [5, 15, 36].

R3: *Enterprise data catalogs should promote a unified understanding of data sets for all user groups by documenting metadata on multiple levels, distinguishing between the conceptual, logical, and physical documentation levels, in order to support heterogeneous user groups in retrieving data.*

Following common metadata standards is also recommended when designing data catalogs. These can be public domain-independent metadata standards or ontologies [8, 15]. Standards promote homogeneous access across heterogeneous descriptions and support data interoperability at the user level [25]. In this way, the utility of data objects is improved, and data consumers and producers are linked by building a common consensus [15, 37]. This influences the interoperability of catalog systems and promotes compliance with FAIR Principles [15]. Concerning the system infrastructure of data catalogs, various metadata standards have already been established, which can be applied in combination depending on the context of use. According to [8], these include the Dublin Core Schema (DC), the Data Catalog Vocabulary (DCAT), the ISO 11179-3 Metadata Registry Metamodel and Basic Attributes (MDR), and the Common Warehouse Metamodel (CWM). Consequently, the requirement is as follows:

R4: *Enterprise data catalogs should support metadata standards to provide users with adequate search results and seamless access to heterogeneous data sets.*

Implementing a business glossary offers advantages for the value and acceptance of the data catalog among users. Clear business terms help to understand the context of the use of the data objects and the data itself by employees of the departments [8, 15, 21, 24]. Business glossaries are central repositories containing key business terms agreed upon by cross-functional subject matter experts [15]. On the one hand, company-wide terms, objects, and attributes can be explained, and on the other hand, domain or business unit-specific terms can be defined [21, 23]. To further optimize the interpretation of the data and their usage environments, the created metadata here can also further be enriched by additional context variables [15]. As a result of a better understanding, the data sets can subsequently be used or adapted for other analysis projects, which is an essential prerequisite for the reusability of the data sets.

R5: *Enterprise data catalogs should be equipped with a complementary business glossary to describe the data objects from an operational perspective to create a uniform understanding regarding specific terms for all user groups and to prevent misinterpretations, given the fact that the user groups come from different domains or companies and have different expertise.*

As integrated platforms that link the various data-oriented user groups (e.g., data owners and data analysts) and enable informal information exchange, it also makes sense to provide efficient data management functions in a centralized manner. These include registration functionalities such as “data connectors” that enable the automatic collection of metadata from source systems or “data imports” that independently import the descriptions of data sets from data tables, which can significantly reduce time-consuming tasks [23]. Furthermore, there are functions for data organization and management (curation of data) that enable, for example, annotations or tags, the creation of metadata, or the labeling of security- and compliance-relevant data [34]. Adding tags or compliance-related information can also influence catalog user collaboration by transparently sharing knowledge and expertise and improving search results. This results in the following requirement:

R6: *Enterprise data catalogs should be equipped with a comprehensive range of data management functions, such as data object registration and curation functions, to facilitate the integration into, the administration of and navigation among the meta data sets.*

Data catalogs are commonly seen as necessary for the implementation of a data governance. This in turn implies that the definition of an enterprise-wide data governance is closely intertwined with the data catalog design. On the one hand, a data governance fosters (or even enforces) compliance with internal and external data management regulations and data protection guidelines and, on the other hand, can support the definition of technical standards to ensure interoperability and thereby maximize data value [22, 23, 39]. In conclusion, data catalogs should fulfill prerequisites that contribute to the implementation of the defined data governance [22]. In this field, the documentation of ownership is an essential prerequisite for assessing responsibilities. This has two benefits. Firstly, contact persons can be identified and contacted directly in case of error occurrences or violations of the defined guidelines. Secondly, contact persons promote collaboration between data consumers and data providers [5, 22]. In addition, knowledge regarding ownership provides information on the relationships between data sets,

allowing important insights to be derived for potential synergies [39]. Thus, ownership representation creates transparency and establishes collaboration opportunities between data consumers and providers. This way, contact persons can be accessed directly in case of questions or problems. In addition, a role model acts as an important prerequisite for system-wide collaboration, as tasks can be distributed and responsible users identified. The following requirement is derived from this:

R7: *Enterprise data catalogs should support clear and consistent data governance structures, including unambiguous role models, ownership, and policies regarding data quality and data provenance that act as an organizational framework to ensure the responsible use and management of data sets.*

Access control mechanisms are central for protecting sensitive data from misuse and complying with regulations [15, 34]. This is true for all data bases but data catalogs in particular which is why their design should include data access functionality. This can include automated workflows for approval processes and user authentication mechanisms [8, 15, 25, 40]. Such functionality ensures that the visibility of catalog content needs to be unlocked by access requests and the assignment of appropriate access keys [5, 41]. As a more recent development, Artificial Intelligence (AI) can be used to identify sensitive or secret data by assigning attributes or to display data sets that are not accessible to the user [15, 23, 24]. Another prerequisite for access control is the definition of user groups and role-specific data authorization levels through which suitable approval processes can be created [21, 23]: Data catalogs should document the approval history and reasons for the access request to analyze the contexts of use of the data and trace potential compliance violations [8].

R8: *Enterprise data catalogs should be equipped with reliable mechanisms for role-specific access controls, secure process flows, and usage policies that regulate data usage, management, and access in terms of security and privacy and that allow only authorized users to access data sets to prevent sensitive data from being misused.*

In addition, data catalogs should ensure the quality and reliability of data and meta-data through various functions. Ideally, the tools encourage the users to define quality standards and measurable data quality metrics in advance and allow to continuously check them later. This way, errors, deviations, and duplicates can be detected early after launching a data catalog [23, 39]. Dashboards can also be a valuable tool for the support of data quality management activities as they can graphically display quality metrics for the selected data sets, visualize developments over time, and signal issues with alerting mechanisms [23, 24]. It should also be possible to add new quality rules or modify existing ones [23]. To ensure the quality of the data in the long term, the users need to continue developing procedures for the maintenance and upkeep of the data sets, including clear responsibilities for each individual process instance. By doing so, it can already be ensured during the context of the design that the catalog system that it can provide coherent and valuable data sets over the entire life cycle of the data catalog [15, 22].

R9: *Enterprise data catalogs should provide adequate control mechanisms in the form of qualitative standards, guidelines, and predefined quantitative data quality metrics that can be continuously reviewed to avoid unreliable or erroneous data objects within the data catalog system.*

Furthermore, there is a need to embed data catalogs in existing infrastructures so that data consumers have standardized access to distributed resource descriptions and information systems [25, 38]. Two building blocks are necessary to ensure sufficient interoperability. Firstly, data catalogs should be equipped with standardized application programming interfaces (APIs) to access the source systems [8, 21, 35, 39]. Of particular interest are interfaces to other data catalogs (especially in large organizations or data ecosystem settings) and the functionality to connect with leading enterprise systems (i.e., ERP, CRM, SCM, CRP, or MES) as well as with business intelligence tools [11]. Secondly, uniform standards, schemas, terminologies, and formal and comprehensively applicable languages for the description of data sets and metadata should be used [15, 24, 25, 37].

R10: *Enterprise data catalogs should incorporate standardized application programming interfaces to query the data sets, their description, and metadata to facilitate the integration into existing technical infrastructures and source systems and give access to different functional units of an organization.*

Since data catalogs should enable both technical and non-technical expert users to access data, user-friendly graphical user interfaces (GUI) are a common essential requirement. Ideally, those GUIs can be parameterized depending on the respective user role [23]. Additionally, data catalogs can include visualization functionalities that advance an understandable and descriptive representation of data sets, metadata, terminology, and data sequences. Data flow diagrams or knowledge graphs have proven to be a viable tool for this [22, 24]. Existing empirical research on data catalog suggests that data analysts value graphical representations of entire metadata collections and logging of historical queries to save users (especially inexperienced ones) the effort to develop queries [16].

In addition, data exploration and visualization tools can be used to display quality metrics or other KPIs in dashboards. They support users in evaluating and analyzing the data [8]. The visualization should enable the various user groups, especially data analysts, to derive insights from the data sets recorded in the data catalog that can contribute to data-related decision-making and the quality assessment and improvement of the data objects.

R11: *Enterprise data catalogs should foster digital interactions of data consumers through intuitive digital user interfaces that meet the needs of non-technical user groups and are thus customizable and allow visualization of data sets.*

Another goal of data catalogs is to promote the collaboration between different data users by providing functions for the exchange of practice-related knowledge and, if necessary, its transfer to other data projects [23]. The progression of transparency regarding the company's existing data objects is crucial to developing a collaborative environment. A characteristic of this is that data sets become traceable and findable for the various user groups [24]. Comment, tagging or rating functions, as well as workflows or discussion forums are useful for promoting communication and collaboration between users of data catalogs [8, 22, 23]. In addition, chat functions can be helpful in establishing direct contact with data owners or contacts and allow clarifying ambiguities or sharing feedback regarding the quality or usefulness of the data [8, 22]. Functionalities for registration, publication, search, filtering, and localization of data sets are additional pillars for a successful data collaboration [35, 42, 43]. In this context, role-specific

functions can be offered that support the fulfillment of the respective tasks and meet the needs of the different user groups [22]. Possible functionalities would be the provision of data preview to gain initial insights into the contents of data sets, the possibility to follow data sets and receive notifications of changes, or recommendations based on previous search queries or user behavior [8, 22, 34]. However, these functions should be provided modularly to offer users only functions that clearly support the specific user role without overstressing the user.

R12: *Enterprise data catalogs should be modularly equipped with collaboration and communication features that enable synergies between data-driven user groups and promote collective decision-making so that users with different levels of knowledge and experience can make better data-based decisions.*

The analysis of the selected publications clearly shows that a high degree of automation is indispensable to achieve the sustainable performance of the data catalog by implementing the previously presented requirements with sufficient performance. There are various use cases for automation in data catalogs, particularly concerning data-driven analysis projects. For example, processes can be automated by incorporating workflows (e.g., approval processes for changes or access requests), or machine learning or artificial intelligence (AI) algorithms can be used in detecting anomalies and causes of errors, analyzing data, or generating insights and recommendations regarding data sets [8, 24]. Furthermore, data description, context enrichment, and metadata generation can be supported using automated approaches. Here, the implementation of machine-based dataset profiling techniques is recommended, with the option to automatically create data profiles [36]. Regarding the principle of “reusability,” an automated documentation of generated analyses results can further be used to derive lessons learned or leverage analysis data for more advanced projects [8]. A nuanced reconstruction of the lineage of data sets can also be recorded in an automated manner, increasing the transparency of the origin of data objects and promoting trustworthiness in the data [23]. The automation dimension indicates that support functions such as AI are needed to facilitate data registration and curation. Furthermore, this has the added benefit that company-wide data catalogs become scalable without losing consistency or accuracy [22, 23, 44]. However, it should be considered that the analytics methods often need to be tailored to the targeted analysis contexts.

R13: *Enterprise data catalogs should be equipped with intelligent automation functions to reduce time-consuming and manual activities of data discovery, analysis, and use on the part of data consumers and time-consuming and manual activities of data management and maintenance on the part of data providers.*

5 Conclusion

Enterprise data catalogs are a “hot topic” in practice to support metadata management. This study elaborates and categorizes a set of 13 functional requirements systematically derived from scientific literature and three practical studies. The main goal of this article is to present a list of relevant functional requirements for practitioners who make decisions on the implementation and tailoring of enterprise data catalogs, to improve their design and increase their acceptance by potential users. The requirements support IT

decision makers in designing and customizing data catalogs to support the integration of data into software-intensive services [3, 4] for the facilitation of software-intensive business operations.

Considering the structure and the priority of these requirements, they cover on a foundational set of base requirements that are crucial for the overall functionality of a data catalog. These are at least partially met by existing open-source or commercial tools. The set of requirements also covers key technical functionalities for data storage, access, and management. Without these, the more user-oriented ones would not work as well, revealing also a natural hierarchy within the requirements set. The different target groups (end users, system operations, database administrators, developers) and their use cases build the foundation for sorting the requirements situationally.

We argue that while our focus originates from an enterprise context, the adoption of data catalogs is also becoming increasingly relevant for non-commercial organizations such as government institutions and nonprofit organizations. In this context, we consider data catalogues as enablers for inter-organizational networks and data ecosystems. This is exemplified in the existing data space or data cooperative initiatives to enable scenarios, such as circular economy, which highly rely on sharing metadata resources at scale [45, 46]. The derived functional requirements are not limited to a particular domain or scenario, and can therefore be used in data-driven scenarios in different domains, although specific tailoring might be necessary. It is also important to consider how the nature of such ecosystems evolves when data catalogues become machine-readable, enhanced by the natural language processing capabilities of current Large Language Models (LLMs). Such advancements enable the connection, processing, and utilization of data in these catalogues with minimal human intervention.

Furthermore, the requirements also help service providers and data catalog solution providers with the integration and customizing of data catalogs. Hence, we are confident that the derived requirements support the value proposition deployment of software companies that offer enterprise data catalogs as software products. Our requirements can also be linked to the Fraunhofer ISST functional model, extending it with prescriptive statements about the functionalities that data catalogs must provide [22]. The requirements can be used for context-specific benchmarks and act as a checklist for system designs or development projects. In addition, the requirements provide a starting point for future design-oriented research on data catalogs. To the best of our knowledge, existing data catalog tools only cover the set of requirements only in a basic manner, especially those focused on end-users (R11-R13). This highlights a significant gap that needs to be addressed.

However, the requirements are mainly limited to the scientific literature, which at this point in time, has done relatively little research on data catalogs. Thus, these results present a synthesized knowledge of the literature but without integration of project experience knowledge from the field. Since domain-specific restrictions (e.g., related to interoperability, standardization or data governance) are not included, the requirements catalog is not exhaustive. Yet, the presented requirements build a foundation for further empirical research on the design of data catalogs capturing domain constraints.

Nevertheless, the requirements catalog should be validated and extended in further studies, especially through empirical cases or the analysis of existing data catalog systems

in order to capture seemingly “trivial” requirements or requirements that reflect the dynamics of the field [14]. The latter is a particular problem given the breathtaking speed at which new AI solutions are introduced to the market which support IT-processes in particular. Therefore, we expect that those reshape the functionality of data catalogs and alter the elicited requirements significantly in the mid-term future. Given R1, it can be assumed that search functionality can be expected to benefit considerably in the near future by applying so called large language models that provide both a more user-friendly natural language interface and can extract semantic similarities. Accordingly, future studies should explore solution approaches for novel AI functions for data catalogs for the new levels of data catalog automation, their effectiveness, shortcomings, and their acceptance. In addition, future research can also explore best practices and strategies for implementing enterprise data catalogs. Ideally, this is done by utilizing the action design research approach in order to combine practical requirements, innovative solutions, and theoretical rigor.

References

1. Legner, C., et al.: Digitalization: opportunity and challenge for the business and information systems engineering community. *Bus. Inf. Syst. Eng.* **59**(4), 301–308 (2017)
2. Dremel, C., Wulf, J., Herterich, M.M., Waizmann, J.-C., Brenner, W.: How AUDI AG established big data analytics in its digital transformation. *MIS Q. Exec.* **16**(2), 81–100 (2017)
3. Hunke, F., Heinz, D., Satzger, G.: Creating customer value from data: foundations and archetypes of analytics-based services. *Electron. Mark.* **32**, 503–521 (2022)
4. Ksouri-Gerwien, C., Ebel, M., Bittner, K., Poepplbuss, J.: Offering knowledge as a service – a taxonomy of knowledge-intensive business services. In: *Proceedings of the 31st European Conference on Information Systems, Kristiansand (2023)*
5. Shanmugam, S., Seshadri, G.: Aspects of data cataloguing for enterprise data platforms. In: *2nd International Conference on Big Data Security on Cloud*, pp. 134–139. IEEE (2016)
6. Otto, B., Jarke, M.: Designing a multi-sided data platform: findings from the International Data Spaces case. *Electron. Mark.* **29**, 561–580 (2020)
7. Gluchowski, P., Gonschorek, E.: Data Catalog – Transparenz durch Dateninventarisierung. *Rethinking. Finance* **3**, 11–14 (2019)
8. Labadie, C.: *Essays on Data Democratization & Protection in the Data-driven Enterprise*. Doctoral thesis, University of Lausanne (2021)
9. Eichler, R., Gröger, C., Hoos, E., Schwarz, H., Mitschang, B.: Data shopping – how an enterprise data marketplace supports data democratization in companies. In: De Weerd, J., Polyvyanyy, A. (eds.) *International Conference on Advanced Information Systems Engineering (CAISE) Forum. LNBP*, vol. 452, pp. 19–26. Springer, Cham (2022)
10. Eichler, R., Giebler, C., Gröger, C., Schwarz, H., Mitschang, B.: Modeling metadata in data lakes – a generic model. *Data Knowl. Eng.* **136**, 101931 (2021)
11. Jahnke, N., Otto, B.: Data catalogs in the enterprise: applications and integration. *Datenbank-Spektrum* **23**, 89–96 (2023)
12. Spezzati, A., Kheradmand, E., Gupta, K., Peras, M., Zaminpeyma, R.: Note: leveraging artificial intelligence to build a data catalog and support research on the sustainable development goals. In: *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*, pp. 579–584 (2022)
13. Dibowski, H., Schmid, S., Svetashova, Y., Henson, C., Tran, T.: Using semantic technologies to manage a data lake: data catalog, provenance and access control. In: *Proceedings of the*

- 13th International Workshop on Scalable Semantic Web Knowledge Base Systems (SSWS 2020), Athens, pp. 65–80 (2020)
14. Zaidi, E., De Simoni, G., Edjlali, R., Duncan, A.D.: Data catalogs are the new black in data management and analytics. Gartner, Consultancy Report (2017)
 15. Ehrlinger, L., Schrott, J., Melichar, M., Kirchmayr, N., Wöß, W.: Data catalogs: a systematic literature review and guidelines to implementation. In: Kotsis, G., et al. (eds.) Database and Expert Systems Applications - DEXA 2021 Workshops. CCIS, vol. 1479, pp. 148–158. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87101-7_15
 16. Gunklach, J., Michalczyk, S., Nadj, M., Maedche, A.: Metadata extraction from user queries for self-service data lake exploration. *Datenbank-Spektrum* **23**, 97–105 (2023)
 17. Altendeitering, M., Guggenberger, T.: Designing data quality tools: findings from an action design research project at Boehringer Ingelheim. In: Proceedings of the 29th ECIS, Marrakesh (2021)
 18. Ehrlinger, L., Wöß, W.: A survey of data quality measurement and monitoring tools. *Frontiers Big Data* **5**, 850611 (2022)
 19. de Reuver, M., Ofe, H., Agahari, W., Abbas, A.E., Zuiderwijk, A.: The openness of data platforms: a research agenda. In: Proceedings of the 1st International Workshop on Data Economy, New York (2022)
 20. Choi, M.-Y., Moon, C.-J., Jung, S.-J.: Building methods of intelligent data catalog based on graph database for data sharing platform. *ICIC Int.* **11**(1), 953–959 (2020)
 21. Mamrot, S., Nowak, F., Ryzyszczak, K., Kaczmarek, Ł., Krzywy, J.: Applying central data catalogues to implement and maintain digital public services. a case study on catalogues of public administration in Poland. In: Janssen, M. et al. (eds.) Electronic Government. LNCS, vol. 13391, pp. 31–46, Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15086-9_3
 22. Jahnke, N., Spiekermann, M., Ramuzeh, B.: Data Catalogs. Implementing Capabilities for Data Curation, Data Enablement and Regulatory Compliance. Fraunhofer Report (2022)
 23. Russom, P.: The Data Catalog’s Role in the Digital Enterprise. TDWI Checklist Report (2017)
 24. Labadie, C., Eurich, M., Legner, C., Fadler, M.: FAIR enough? Enhancing the usage of enterprise data with data catalogs. In: Proceedings of the 22nd Conference on Business Informatics (CBI), pp. 201–210. IEEE (2020)
 25. Quimbert, E., Jeffery, K., Martens, C., Martin, P., Zhao, Z.: Data cataloguing. In: Zhao, Z., Hellström, M. (eds.) Towards Interoperable Research Infrastructures for Environmental and Earth Sciences. LNCS, vol. 12003, pp. 140–161. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-52829-4_8
 26. Kerhervé, B., Gerbé, O.: Models for metadata or metamodels for data? In: Proceedings of 2nd IEEE Metadata Conference, Silver Spring, pp. 1–12 (1997)
 27. Riley, J.: Understanding Metadata. What is metadata and what is it for? <https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf>. Accessed 26 Feb 2023
 28. Oram, A.: *Managing the Data Lake*. O’Reilly (2015)
 29. Diamantini, C., Giudice, P.L., Musarella, L., Potena, D., Storti, E., Ursino, D.: A new metadata model to uniformly handle heterogeneous data lake sources. In: Proceedings of the 22nd European Conference on Advances in Databases and Information Systems (ADBIS 2018), pp. 165–177 (2018)
 30. Research Data Alliance Homepage. <https://www.rd-alliance.org/groups/metadata-ig.html>. Accessed 26 Feb 2023
 31. Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *Br. J. Manag.* **14**(3), 207–222 (2003)

32. Garousi, V., Felderer, M., Mäntylä, M.V.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Inf. Softw. Technol.* **106**, 101–121 (2019)
33. Mayring, P.: *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, Beltz (2015)
34. Wells, D.: *The Ultimate Guide to Data Catalogs*. White Paper of the Eckerson Group (2018)
35. Lapi, E., Tcholtchev, N., Bassbouss, L., Marienfeld, F., Schieferdecker, I.: Identification and utilization of components for a linked open data platform. In: *IEEE 36th Annual Computer Software and Applications Conference Workshops*, Izmir (2012)
36. Skopal, T., Klimek, J., Necasky, M.: Improving findability of open data beyond data catalogs. In: *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*, pp. 413–417 (2019)
37. Barbosa, E.B., Sena, G.: Scientific data dissemination a data catalogue to assist research organizations. *Ciência da Informação* **37**(1), 19–25 (2008)
38. Stillerman, J., Fredian, T., Greenwald, M., Manduchi, G.: Data catalog project—a browsable, searchable, metaIndata system. *Fusion Eng. Des.* **112**, 995–998 (2016)
39. Joshi, D., Pratik, S., Rao, M.P.: Data Governance in Data Mesh Infrastructures: The Saxo Bank Case Study. In: *Proceedings of the International Conference on Electronic Business*, Nanjing (2021)
40. Lefebvre, H., Legner, C., Fadler, M.: Data democratization: toward a deeper understanding. In: *Proceedings of the 42nd International Conference on Information Systems*, Austin (2021)
41. Czajkowski, K., Kesselman, C., Schuler, R.: ERMrest: a collaborative data catalog with fine grain access control. In: *13th International IEEE Conference on e-Science*, Auckland (2017)
42. Shi, C., Zhang, Y., He, R.: Design and implementation of a P2P resource sharing system based on metadata catalog. In: *Proceedings of the 9th International Symposium on Computational Intelligence and Design*, Hangzhou (2016)
43. Holl, P., Gossling, K.: Midas: towards an interactive data catalog. In: Gadepally, V., et al. (eds.) *Heterogeneous Data Management, Polystores, and Analytics for Healthcare*. LNCS, vol. 11721, pp. 128–138. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33752-0_9
44. Labadie, C., Fadler, M., Eurich, M., Legner, C.: All hands on data: a reference model for enterprise data catalogs. In: *Essays on Data Democratization & Protection in the Data-Driven Enterprise*, pp. 71–108 (2021)
45. Serna-Guerrero, R., Ikonen, S., Kallela, O., Hakanen, E.: Overcoming data gaps for an efficient circular economy: a case study on the battery materials ecosystem. *J. Cleaner Prod.* **374**, 133984 (2022)
46. Jäger-Roschko, M., Petersen, M.: Advancing the circular economy through information sharing: a systematic literature review. *J. Cleaner Prod.* **369**, 133210 (2022)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Are Business Expectations Aligned with the Development Plan Made by the Software Architecture Area? A Case Study on Agile Teams in a Large Company

Marcelo Augusto da Silva¹ , Inaldo Capistrano Costa¹ ,
and Eduardo Martins Guerra² 

¹ Instituto Tecnológico de Aeronáutica - ITA, São José dos Campos, SP, Brazil
marcelo.augusto@ga.ita.br

² Free University of Bozen-Bolzano, Bolzano, Italy

Abstract. In the current scenario of digital transformation, understanding the interaction between the areas of business and software architecture is essential for delivering successful projects. This research aims to elucidate perceptions related to both domains, thus seeking a more efficient collaboration in the context of agile software development projects. Based on a qualitative research method, we conducted semi-structured interviews with product owners and software architects. The collected data were analyzed using Thematic Analysis to discover patterns and themes regarding the perceptions of the interviewed professionals. We found out that business areas often have a limited understanding of the technical complexities involved in software architecture, while software architects sometimes have no knowledge about business development plans. However, a continuous iteration process, supported by proper communication channels, could drive better project results. The study also revealed the potential for a proactive, integrated approach to architecture, focusing on continuous education and team alignment. Finally, bridging the knowledge gap and fostering collaboration between the two areas may lead to more efficient and effective software development processes. Future research perspectives could reveal strategies that would improve this collaboration or explore similar dynamics in different organizational contexts.

Keywords: Agility · Software Architecture · Agile Methodology · Case Study · Thematic Analysis

1 Introduction

In the current scenario of software project development, the need for quick delivery and the ability to adapt to constant changes in business requirements are key factors to deliver successful projects. The adoption of agile methodologies

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 19–34, 2024.

https://doi.org/10.1007/978-3-031-53227-6_2

emerges as a response to this demand, thus promoting greater flexibility, collaboration, and continuous delivery of value [5]. However, while agile methodologies focus on adaptability and customer interaction, software architecture remains a complex technical aspect that can often be overlooked. This dichotomy between agile adaptability and the need for a solid architecture may lead to misalignments between the expectations of the business area and the software product that has been developed [4].

The alignment between the expectations of the business area and the software development project team is essential to ensure that the technological solution that has been developed is in sync with the business goals [12]. This approach not only enhances the chances of meeting the referred business requirements, but also ensures that the software is efficient, scalable, and sustainable in the long term [13]. Integrating the architecture team into the development process is essential to guarantee that the software can evolve along with the ever-changing demands of the business [6]. Through effective collaboration, the understanding of business objectives becomes clear, and the software architecture team can provide the necessary guidelines for a successful implementation [14]. The lack of such alignment may result in solutions that fail to meet the needs of the business, in addition to presenting technical challenges, thus affecting system availability, performance, and maintenance [15].

This article aims to explore these potential misalignments, taking as a case study the software project development environment of a large cooperative financial system in Brazil - the software developed by the referred organization employs agile development practices and is widely used throughout the country. Thus, through this research, we aim to understand the nature of these eventually existing discrepancies and offer insights that can help development teams harmonize agile practices with the requirements demanded by software architecture, thus ensuring that both walk side by side in favor of more aligned and effective solutions. To investigate this phenomenon, we were guided by the following research questions:

- RQ1: How does the business area perceive software architecture and what relevance do they assign to the development of software projects?
- RQ2: What is the level of knowledge of the software architecture area in relation to the application development plan?
- RQ3: How do the architecture and business teams perceive the iterations in the development process of software projects?

In order to answer our above-mentioned research questions, we conducted ten semi-structured interviews [1] with professionals who are currently working on software development projects in the environment of the referred financial cooperative in Brazil. The study included five professionals who are currently working in the business area and are responsible for stating the requirements that the software must meet and five other professionals who work in the technology area and are responsible for structuring the architecture that the software must follow to be implemented. We then proceeded to an inductive thematic analysis [2,3] of the interview transcripts.

In this study, the results emerge as a deep reflection of the existing dynamics between the areas of business and software architecture in contemporary organizations. Throughout the analysis, we were able to reveal distinct and sometimes conflicting insights about the role and relevance of software architecture in the context of project development. Such findings throw light on areas of misalignment and also identify potential for optimization in the collaborative process between technical and business teams, thus suggesting an intrinsic need for realignment in order to deliver more effective software solutions.

The contributions of this work go beyond the mere identification of these dynamics, providing a practical road map to facilitate effective integration between architecture and business teams. Based on the recommendations proposed herein, this study acts as a guide for organizations seeking to strengthen their collaborative approach, emphasizing the importance of mutual understanding and aligned goals. The insights and strategies presented in this paper can potentially serve as a reference point for organizations willing to align their technical initiatives with their business strategies more effectively.

The article is structured as follows: Sect. 2 presents the Software Architecture theme and its relevance in software projects. Section 3 contextualizes our research by connecting it to similar studies on the subject. Section 4 details the method and tools employed in our data collection and analysis. In Sect. 5 we present and discuss what we found by analyzing the interactions between the referred areas during project development. We come to a conclusion in Sect. 6, where we reflect on our findings and point to possible directions for future research.

2 Software Architecture Relevance

Software architecture can be understood as the structure of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution. It establishes the fundamental organization of a system in terms of its components and their interactions, and is critical to determining software quality, performance, and longevity. The IEEE, in its standard definition, describes software architecture as “the fundamental structure of a system, which consists of software components, their externally visible properties, and the relationships among them” [29].

Bass et al. [6] describe software architecture as the structure of a system that includes software components, the relationship between these components, and the properties of both elements. In this context, software architecture is more than just the structure; it also defines how the components interact and how the structure evolves over time.

According to Shaw and Garlan [21], software architecture is a discipline that provides a structural point of view and provides techniques to help create highly structured and modular systems.

The interaction between software architecture and non-functional requirements (NFRs) plays a crucial role in the software project development process.

NFRs, such as performance, security, and reliability, significantly influence architectural decisions. Certain quality attributes hold pivotal importance in architectural design stages, due to their direct impact on the system’s structure and design pattern choices. Additionally, proper categorization of NFRs is necessary for effectively evaluating software architecture. Simultaneously, managing these requirements in specific development contexts, like model-driven development, underscores the need for a systematic approach from the outset. This integration fosters better conditions for the final architecture to align with stakeholders’ expectations and the system’s operational requirements [6,32,33].

In parallel, agile software development processes have gained prominence due to their ability to provide value in an iterative and incremental way, prioritizing collaboration and response to change. However, aligning architecture strictness with the flexibility found in agile methods can be a challenge. Architectural decisions often require early planning and consideration, while agile methods value adaptation and continuous delivery. Thus, to achieve optimal balance, it is vital that development and architecture teams collaborate closely and adjust their processes and practices in order to align the benefits of robust architecture with the agility of development processes [6].

In the organization studied, software architecture plays a leading role, which is evidenced by the existence of a unit in the IT sector consisting of professionals specialized in this field with the purpose of satisfying the inherent needs of software development projects. This unit actively collaborates with sectors vital to IT such as security, infrastructure, and operations, so that solutions reach adequate standards of security, availability, and robustness. This organization generates solutions at a national scale, serving approximately 8 million users who carry out financial transactions both in person at business units and via self-service channels. It is also important to mention that the organization operates in a highly regulated sector of the economy. Thus, its software projects are often shaped by external influences, which include transactions that must adhere to SLAs determined by regulatory bodies, for example.

In summary, software architecture provides a blueprint for the system, representing its main properties and how they interact. It is the key artifact for understanding any system’s large components and how they are orchestrated to work together.

3 Related Works

Upon investigating the existing literature on the alignment between the business and the software architecture areas as well as the impact of organizational models on agile development, several prominent works were identified. These works provide a critical perspective on the challenges, solutions and trends associated with this subject. Within the context of the research questions included in this study, we can highlight the following works.

Rozanski and Woods [7] delve deeply into software architecture and the relationship with stakeholders; they don’t specifically focus on “the business perception of software architecture” as an isolated topic. Instead, they provide a

comprehensive approach to address the concerns of all stakeholders, including but not limited to the business itself. The main focus of this work is to provide a structured approach to software architecture and communicate this architecture to stakeholders.

Garlan [20] in turn, discussed how software architectures often evolve in response to external pressures. Market changes, new technologies, and the emergence of competing standards may lead to unplanned adjustments in architecture. This study highlights the importance of a flexible and adaptable architecture to address these challenges.

Research by Dingsøyr et al. [8] highlighted that continuous collaboration and frequent iterations are essential for agile development. They noticed that teams that work closely together and review their processes regularly are more likely to understand and implement requirements effectively, which results in higher-quality software.

Kniberg and Ivarsson [18], in their famous white paper about the Spotify model, described how guilds and other organizational structures can promote collaboration and knowledge sharing. Their work provides robust evidence that such frameworks can mitigate challenges that are commonly faced in software development, especially those related to communication between technical and business areas.

The study by Viviani et al. [31] highlights the critical management of NFRs in software projects, emphasizing their propensity for change and late definition, aspects often underestimated in software architecture planning. The research, through responses from professionals with extensive experience, revealed that NFRs undergo significant alterations, often late in the development cycle, highlighting a notable gap in the elicitation, validation, and management of these requirements. This discovery underscores the pressing need for agile approaches that can accommodate such uncertainties and changes, ensuring that the software architecture maintains its integrity and relevance over time, considering that the change and evolution of NFRs are inevitable in the software evolution cycle.

The above-mentioned works highlight the complexity and importance of effective alignment between the business area and the technical teams. Proper integration and continuous communication are essential to ensure that the software developed is aligned with the company's goals and needs.

4 Research Method

To deepen the understanding of misalignments between the business area's expectations and the architectural solutions implemented in software projects within the studied organization, a case study approach was chosen [17] with a qualitative research method. Semi-structured interviews were used as the data collection instrument [2,3] with professionals involved in the software development process.

Interviews. From May 2022 to July 2023, ten interviews were conducted with professionals who work on software projects in the organization studied. Initially, two interviews were carried out: one with a software architect and the other with a product owner. A preliminary analysis of these data was carried out to determine whether they would be adequate to guide our research development. After this initial assessment, the interviews continued. All sessions were conducted online in Portuguese through video conference and lasted about 45 min each.

Participants. In order to assess the perception of professionals in the business area and those responsible for software architecture, five professionals corresponding to each profile were selected. The business professionals interviewed were appointed by managers of business product areas and the IT professionals - all software architects - were appointed by the manager responsible for the software architecture area. After being assigned to participate in the research by their respective managers, all were duly contacted, briefed on the issue under study, and invited to voluntarily participate in the research. All the appointed professionals agreed to participate and therefore the interview session was scheduled. Figure 1 details the interviewees' qualifications. In the interview session, which was recorded with prior authorization from the participants, previously prepared questions were presented to the participants who then expressed their perception about the issue raised.

Profile	Quantity	Gender		Experience (Year)			
		Male	Female	<= 2	> 2 <= 4	> 4 <= 6	> 6
Software Architect	5	5	0	0	0	3	2
Product Owner	5	2	3	2	1	0	2

Fig. 1. Profile of the interviewees.

Research Ethics. During the recruitment process, participants were informed about the purpose of the study, the content of the questions, and the affiliation of the interviewer. In the organization studied, it is widely known that there are professionals on their staff, people who take up a professional master's degree course which is encouraged by the organization itself. Aware of this condition, participants agreed to participate in this study which can bring benefits to the organization's software development process. At the beginning of each interview, the interviewer made sure to announce the purpose of the study and the anonymous nature of its content, in addition to explaining the dynamics of the interview and obtaining verbal consent from the interviewee. Since the interviews were conducted using Microsoft Teams¹, they were recorded with the participant's consent and transcribed automatically by the tool itself during the course of the interview, and the interviewee also viewed the content of the transcript.

¹ <https://www.microsoft.com/pt-br/microsoft-teams/log-in>.

Data Analysis. To analyze these transcripts, we developed an inductive coding scheme. Inductive coding was used to investigate the participants' insight into the software project development process in the organization studied - the aim was to identify dysfunctions that would create a gap between what the user expects the software to deliver throughout its development and how the project development area prepares the software architecture to meet present and future requirements. In this approach, themes emerge from the data, and codes are signed when concepts become apparent in these data. This means that the researcher encodes the data without trying to fit them into a pre-existing coding framework or their own analytical biases [2].

To develop the analyses, the ATLAS.ti² software was used and the thematic synthesis process proposed by Braun and Clarke [2] was followed. A researcher began the analysis by carefully reading the transcripts and getting immersed in the data. Subsequently, specific text segments were identified, labeled, and transformed into initial codes. To ensure coding accuracy and cohesion, a random selection of these codes was submitted to the research group for evaluation. This allowed for a uniform understanding of the codes among the members. The following step involved the conversion of these codes into themes, which were subdivided into sub-themes and higher-order themes. The researcher then thoroughly reviewed all themes and data, ensuring their congruence, which led to the elaboration of a thematic map of the analysis. To add rigor to the process, another researcher was introduced to reassess the codified texts and established themes. The final structure of themes and sub-themes emerging from the analysis can be viewed in Fig. 2 - details and further discussion will be covered in the subsequent section.

5 Results and Discussion

Upon carrying out semi-structured interviews, it was possible to identify key patterns and themes related to the interaction dynamics between the business, development, and software architecture teams in the context of project development. The main findings have been organized into 5 themes, as follows:

5.1 Established Architectural Infrastructure

One of the main findings of this study refers to the existence of a well-established reference architecture in the organization which, in general terms, is aligned with the non-functional requirements of the various software developed and used in the referred environment. This implies the existence of a pre-defined set of standards, principles, and components that are considered standard for the construction and evolution of systems. Reference architecture serves as a blueprint, ensuring that systems are consistent, interoperable, and aligned with organizational strategy. One of the interviewed architects made the following statement:

² <https://atlasti.com/>.

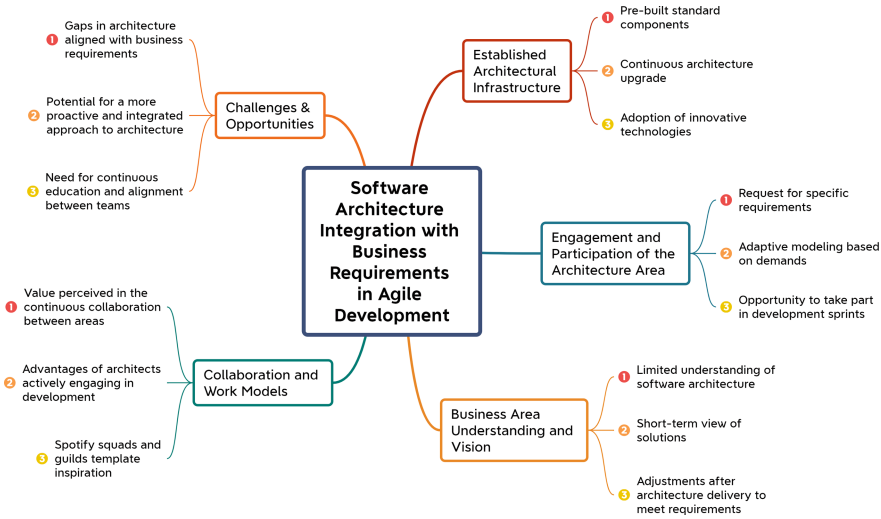


Fig. 2. Final thematic map.

“reference architecture would be a guide to good practices. Practices that must be adopted as a norm. As a rule, they are drivers to be applied to business data” [Architect-1].

We must emphasize the importance of reference architectures, as they provide a solid foundation for development and help reduce costs by avoiding rework and speeding up delivery by reusing previously validated components [6]. This architecture may also help ensure regulatory or security standard compliance, in addition to facilitating communication between teams as it creates a common language and shared understanding regarding standard technical solutions [7].

We also identified that this reference architecture has a regular evolution plan that seeks to provide modern solutions, compatible with what is offered by the market, thus keeping the software ready to meet the referred business requirements. Among the various reports on the maintenance of this reference architecture, one of the architects stated: “We have an overall plan when it comes to creating new components, not a specific architecture plan. There’s the creation of new products and everything must be done in a new architectural design” [Architect-5]. Shaw and Garlan [21] stated that as business needs and the technological scenario evolve, software architecture must be adjusted and reviewed to continue meeting emerging requirements and challenges.

Another relevant matter about architecture maintenance identified in this study refers to prospecting innovative technologies, as mentioned by one of the interviewed architects, “Among its attributions, the architecture team must prospect new technologies and bring them to the company and, in a way, make them operational, so that these technologies can be used by the development teams” [Architect-4]. Foote and Yoder [22] mention that evolution and innovation are inseparable in the context of software development. By introducing

innovative technologies, it is possible to address new challenges and optimize the systems' performance and efficiency.

Finally, we were able to verify that the architects' statements showed that they are committed to promoting continuous evolution, adopting good practices, and incorporating innovations, which may be an indication of the organization's architectural maturity.

5.2 Engagement and Participation of the Architecture Area

Based on the statements given by the interviewees, it was possible to identify a series of practices and challenges related to the engagement of the architecture area in the development of software projects in the organization featured in this study. These observations are in line with existing discussions in the literature about the role of architecture in agile teams and the integration of architects in development teams.

The participation of architects often begins when there are specific demands that require technical assessments. This can be evidenced in the statement given by one of the interviewed architects: *"We need to assess if [this demand] will have support. So what shall we do? Shall we have a chat about it? Let's call an alignment meeting to discuss some architectural pre-documentation aspects related to software architecture"* [Architect-1]. Figure 3 shows a representation of the identified working model. Kruchten [23] argues that, in agile environments, teams often find themselves in situations where architectural expertise becomes vital, especially when new challenges arise.

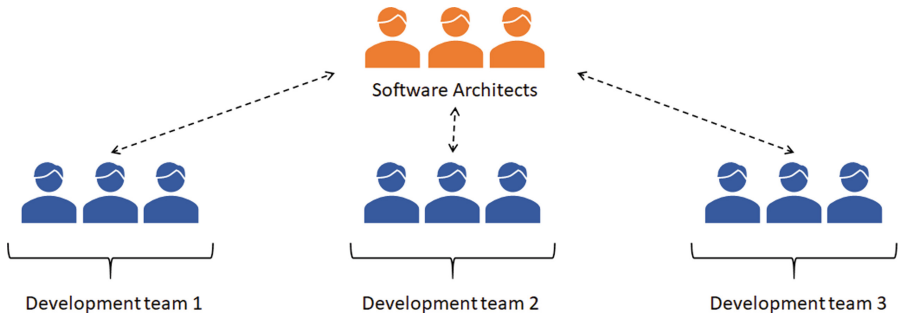


Fig. 3. A Separate Team of Software Architects Works with Multiple Development Teams [30].

Another architect made the following comment: *"a call comes in for us to assess some data, for example, and that's when we become aware of what is being developed"* [Architect-1]. Currently, architects are called upon mainly when specific architectural demands arise. This model may result in late design decisions and possible rework if architectural considerations are not duly identified at the beginning of the development cycle [6].

As for an earlier performance of architects along with the development teams, one of the interviewees made the following comment: *“The software architecture’s pre-documentation meeting does not exist. It all comes from the gut feeling of the development team, really. So I would say that the software architecture documentation is the trigger for us to start being aware and acting together with the team, but this informal approach with teams that have more expertise, as I mentioned before, may occur. This is what we must assess”* [Architect-1]. A proactive participation of architects throughout the whole development cycle may facilitate the early identification of architectural challenges, allowing for solutions that are more knowledgeable and aligned with business needs and technical constraints [9]. In addition, their constant presence may serve as an ongoing education channel for the business team, helping them understand the role and importance of software architecture in their projects.

In summary, the interaction between architects and development teams, as observed in the above-mentioned statements, points to a need for greater integration and continuous collaboration. Practices observed in the organization featured in this study, although in line with some published works, also suggest that there are opportunities for a more systematic and continuous approach to architectural engagement. Encouraging closer collaboration between business, development, and architecture areas can not only improve the quality of the delivered solutions but also promote a more harmonious working environment, with fewer conflicts and misunderstandings [10].

5.3 Business Area’s Understanding and Views

In the agile development environment, the role of software architecture is often underestimated or misunderstood, especially by business teams [11]. Evidence of that could be identified in the organization studied, where the business team has little clarity on what constitutes software architecture and how relevant this component is to project delivery. This fact was evidenced by the speech of one of the professionals in the business area, as highlighted below, about what software architecture is: *“I’ll tell you my understanding of it based on the little contact I’ve had. I understand that they create a framework and from that framework, they can build something there. I don’t know exactly what that is”* [Product Owner-5].

Another feature identified in the study is the absence of a structured, long-term product development plan. In terms of evolution and innovation, responses mentioned incremental deliveries. One of the respondents said: *“On the product itself, I don’t see much change for the next 5 years, as a form of business. It basically depends on the Central Bank”* [Product Owner-4]. The tendency to focus on immediate needs and not anticipate changes over a long-term horizon may lead to decisions that are not scalable or flexible [24]. The regulatory role of the Central Bank, as mentioned by the interviewee, also highlights the importance of considering external factors that may influence product decisions and their development.

The observation that architectural adjustment often occurs in response to significant incidents, as mentioned by one of the respondents - *“It usually comes*

after an incident happens. After there's been a lot of fuss over it..." [Architect-5] - highlights an often delayed response to changing needs. This reactive type of approach may lead to one-off solutions and possibly more costly and complex refactoring operations in the future.

Guerra et al., in their study, present the idea of "architectural triggers", which are predefined events or conditions that indicate the need for architectural reviews or adjustments [25]. These triggers can be seen as a proactive approach, allowing teams to identify and respond to potential architectural issues before they evolve into significant crises. By incorporating such triggers into the development process, professionals can better anticipate and manage necessary changes, ensuring that the software develops in a more controlled and sustainable manner.

The above-mentioned observations suggest that there is a need for better alignment and communication between the business and technical areas, ensuring that the long-term implications of architectural decisions are well understood and taken into consideration in the development of software projects.

5.4 Collaboration and Work Models

Collaboration and effective communication between different areas of software engineering are critical to ensure that end products are robust, scalable, and meet end-user needs. This collaboration is essential not only between individual members of the software team but also between different teams and areas of expertise [6]. The comment made by one of the interviewees - *"Not just architecture, but also infrastructure and tests should be part of my day-to-day business development. Without silos. Today it's not like that here. I think this [approach] makes things very complicated"* [Architect-1] - on the need for a daily collaboration reflects the opinion of many authors who suggest that integrated and collaborative teams are more effective in delivering high-quality software [26].

Furthermore, the emergence and success of multidisciplinary team models, as highlighted by another interviewee - *"Why did they decide to break up and create these cross-functional teams? Because now the success metric of that whole team, that cross-functional team, happens to be the project, which in the end is what matters to the client"* [Architect-4] - resonate with the advantages perceived in agile development and continuous integration. Figure 4 shows an image that represents the above-mentioned work model. The agile method has become one of the most adopted software development methodologies precisely because it focuses on collaboration, continuous feedback, and adaptation to change [27].

Spotify's model of squads and guilds, also mentioned by one of the interviewees - *"in addition to chapters, you can use the collaboration of other organizations, like, you can organize the chapters, which is done between teams, but you can also have tribes that are larger groups working on a similar business pillar. So, there are still other organizations you can use to redeploy teams"* [Architect-4] - is a particularly successful adaptation of Agile and Lean principles. Not only does it bring together multidisciplinary teams (or "squads") that have autonomy and responsibility for delivery, but it also allows for effective

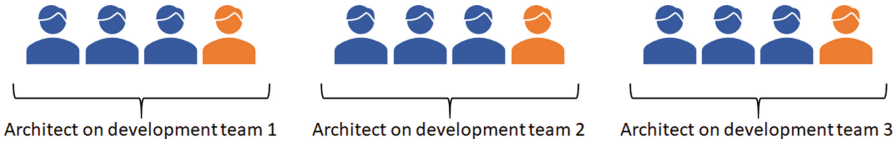


Fig. 4. Each Development Team Has One Software Architect [30].

cross-communication through “guilds”, ensuring that knowledge is shared across squads and that there is consistency where necessary [19].

Nevertheless, it is worth noting that while models such as Spotify’s may work well for some organizations, a successful implementation of these models will depend on the organization’s culture, structure, and goals. Thus, it is essential that companies, such as the one studied in this case, carefully consider their individual needs and contexts before adapting such practices [8].

5.5 Challenges and Opportunities

The role of software architecture in IT projects is crucial not only in terms of technical decisions but also to ensure that the final solution is aligned with the needs of the referred business. However, comments made by the interviewees have highlighted some substantial challenges that must be assessed and addressed.

One of the product owners addressed a common problem in software projects - *“we end up finding it a bit hard to get the expected result. Sometimes when we are talking to the professionals responsible for the requirement analysis process or when we speak to the designer, we create [the product] in some way, and then, when it comes to developing it, we end up finding many barriers in this regard”* [Product Owner-2] - in which there is a disconnect between the initial requirements, the proposed design and the actual deployment [28]. This often leads to rework, project delays, and solutions that do not fully meet the expectations or needs of the business. This lack of alignment emphasizes the importance of clear and effective communication during every stage of the project, as well as the need for a flexible architecture that can adapt to changes as they arise [6].

The observation made by another product owner who participated in the interview suggests a need for greater integration and collaboration between software architects and development teams - *“We spend a lot of time thinking about things like: Is this how we are going to create it? Shall we do it like this? No, but it doesn’t have to be like this, or sometimes even developers have some ideas that can make our processes a lot easier. It goes back and forth, multiple times because I feel like there is this gap”* [Product Owner-2]. As noted by Fairbanks [11], a more proactive approach to architecture can lead to more robust and effective solutions. Furthermore, this collaboration may serve as a means of sharing knowledge and best practices, thus ensuring that everyone on the team is on the same page.

Finally, the comment - *“Teams, they work in different ways, which ends up being a complicating factor when we have to deal with several products. We end*

up having to use different approaches with different teams and that ends up complicating our day-to-day work” [Product Owner-1] - made by another product owner who participated in the interview highlights the challenge of working with multiple teams that may have different methodologies or work patterns. Dingsøyr et al. [8] noted that the standardization of work methods may improve the efficiency and quality of the developed software. However, it is essential to recognize and respect differences between teams and find a balanced approach that allows for flexibility while maintaining a certain level of standardization.

5.6 Validity Discussion

Case studies, especially within the context of software engineering research, face several validity challenges. Runeson and Höst [17] outline various threats to validity in case studies, and these can be extrapolated and applied to the case study carried out in this research, which used qualitative research with thematic analysis. Examining the validity of our study in terms of internal, external, construct validity and reliability, we offer the following considerations:

Internal validity is typically linked to studies aiming to establish causal relationships and elucidate specific conditions or problems [17]. Since our research sought to understand misalignments in software development without emphasizing causal connections, we did not dwell on internal validity. External validity, in turn, assesses whether findings can be generalized beyond the studied contexts. Our results come from a major Brazilian financial institution. To expand the generalization of these insights, additional research in different industries or regions based on more extensive samples is recommended.

Construct validity refers to the alignment of collected data with research questions. In this scope, we prepared a questionnaire and tested it with a product owner and a software architect. Subsequent analysis ensured that the data properly covered the research topic. Moreover, we interviewed experienced professionals from the studied organization, who were appointed by their respective managers.

Lastly, reliability is linked to the objectivity of data analysis, regardless of the researchers involved. At this stage, the first researcher established a case study protocol to ensure consistency in the research methodology. The analysis of the collected data was conducted by the first and second researchers, aiming to ensure a comprehensive view of the software development process and its potential to build an architecture that meets business demands. Consequently, a third researcher reviewed the classifications carried out by the other researchers in order to give them more objectivity and impartiality.

6 Conclusion

The present study aimed to deepen the understanding of the relationship between the business area and software architecture in the context of project

development. Through our investigations, we managed to identify that the perception of the business area on software architecture is diverse - many see it as a fundamental structure to build or adapt functionalities, while others have a more limited view, focused on immediate deliveries (RQ1). This perception may vary, but it reinforces the importance of clear and continuous communication between teams in order to guarantee the effectiveness of the project.

When it comes to understanding software architecture in terms of application evolution, we can see that there is an effort to keep up-to-date and aligned with demands and changes in the business plan. However, it is challenging to keep in sync, given the dynamic nature of businesses and the rapid changes in technology (RQ2).

The perception of iterations in the development process revealed the need for closer collaboration between the architecture and business teams. The introduction of agile practices and collaborative models, such as those inspired by Spotify, may be a promising way to improve this integration (RQ3). However, continuous alignment and the formation of cross-functional teams are essential factors to overcome the challenges identified in this study.

Finally, our study found that the software development process in the studied organization is more exposed to misalignments between the expectations of the business area and the developed solution. However, this fact does not imply that the results of the delivered projects fail to meet expectations. Such a phenomenon was not studied in this research. Still, it is important to point out that the development and deployment processes, when not optimized, can lead to multiple iterations, potential rework, and late delivery.

Our findings highlight the importance of mutual understanding between business and software architecture areas, revealing knowledge gaps and friction points in the context of development process iterations. By elucidating these challenges, the study offers insights for organizations to seek closer and more integrated collaboration, thus promoting greater efficiency in project development. This improved understanding may encourage targeted training, adjustments in organizational models, and the introduction of appropriate collaboration tools, thereby leading to heightened performance in both sectors.

We believe that the findings presented in this study may serve as a starting point for future investigations and improvements in the field of software engineering. In future studies, it may be beneficial to delve into practical strategies to improve communication between the areas of business and architecture, explore the impact of different organizational models on effective collaboration, and investigate how tools and technology platforms can be used to facilitate mutual understanding. Additionally, it would be of great value to analyze the evolution of these interactions over time, considering the rapid changes in technology and business demands, as well as to deepen studies on continuous education and alignment mechanisms between teams in agile environments.

References

1. Adams, A., Lunt, P., Cairns, P.: A qualitative approach to HCI research (2008). https://oro.open.ac.uk/11911/1/9780521870122c07_p138-157.pdf. Accessed 31 Aug 2023
2. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qual. Res. Psychol.* **3**(2), 77–101 (2006). <https://doi.org/10.1191/1478088706qp063oa>
3. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: *International Symposium on Empirical Software Engineering and Measurement*, pp. 275–284. IEEE (2011). <https://doi.org/10.1109/esem.2011.36>
4. Stal, M.: Refactoring software architectures. In: *Agile Software Architecture [S.l.]*. Elsevier (2014). <https://doi.org/10.1016/B978-0-12-407772-0.00003-4>
5. Babar, M.A., Brown, A.W., Mistrík, I. (eds.): *Agile Software Architecture: Aligning Agile Processes and Software Architectures*. Newnes (2013)
6. Bass, L., Clements, P., Kazman, R.: *Software Architecture in Practice*. Addison-Wesley (2012)
7. Rozanski, N., Woods, E.: *Software Systems Architecture: Working with Stakeholders Using Viewpoints and Perspectives*. Addison-Wesley (2012)
8. Dingsøy, T., Nerur, S., Balijepally, V., Moe, N.B.: A decade of agile methodologies: towards explaining agile software development. *J. Syst. Softw.* **85**(6), 1213–1221 (2012). <https://www.sciencedirect.com/science/article/pii/S0164121212000532>. Accessed 31 Aug 2023
9. Kruchten, P.: Contextualizing agile software development. *J. Softw. Evol. Process* **25**(4), 351–361 (2013). <https://doi.org/10.1002/smr.572>
10. Ambler, S.W.: *Agile architecture strategies for scaling agile development*. Agile Model. (2012)
11. Fairbanks, G.: *Just Enough Software Architecture: A Risk-Driven Approach*. Marshall Brainerd (2010)
12. Laudon, K.C., Laudon, J.P.: *Management Information Systems: Managing the Digital Firm* (2004)
13. Sommerville, I.: *IEEE software and professional development* (2016). <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7420501>. Accessed 31 Aug 2023
14. Fowler, M.: *Patterns of Enterprise Application Architecture* (2002)
15. Kruchten, P.: *The Rational Unified Process: An Introduction* (2003)
16. Yin, R.K.: *Case Study Research: Design and Methods*. Sage Publications (2013)
17. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empir. Softw. Eng.* (2009)
18. Kniberg, H., Ivarsson, A.: *Scaling Agile @ Spotify with Tribes, Squads, Chapters and Guilds* (2012). <https://blog.crisp.se/wp-content/uploads/2012/11/SpotifyScaling.pdf>. Accessed 31 Aug 2023
19. Smite, D., Moe, N.B., Levinta, G., Floryan, M.: Spotify guilds: how to succeed with knowledge sharing in large-scale agile organizations. *IEEE Softw.* **36**(2), 51–57 (2019). <https://doi.org/10.1109/MS.2018.2886178>
20. Garlan, D.: Software architecture: a roadmap. In: *The Future of Software Engineering*, pp. 91–101 (2010). <https://doi.org/10.1145/336512.336537>
21. Shaw, M., Garlan, D.: *Software Architecture: Perspectives on an Emerging Discipline*. Prentice Hall (1996)
22. Foote, B., Yoder, J.: The selfish class. In: *Proceedings of the 3rd Conference on Pattern Languages of Programming* (1998)

23. Kruchten, P.: Agility and architecture: can they coexist? *IEEE Softw.* **27**(2), 16–22 (2010). <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5420791>. Accessed 31 Aug 2023
24. Boehm, B., Turner, R.: *Balancing Agility and Discipline: A Guide for the Perplexed*. Addison-Wesley (2004)
25. Wirfs-Brock, R., Yoder, J., Guerra, E.: Patterns to develop and evolve architecture during an agile software project. In: *Proceedings of the 22nd Conference on Pattern Languages of Programs*, pp. 1–18, October 2015. <https://doi.org/10.1145/3424771.3424800>
26. Cohen, D., Lindvall, M., Costa, P.: An introduction to agile methods. *Adv. Comput.* **62**, 1–66 (2004)
27. Highsmith, J., Cockburn, A.: Agile software development: the business of innovation. *Computer* **34**(9), 120–127 (2001)
28. Boehm, B., Basili, V.R.: Software defect reduction top 10 list. *Computer* **34**(1), 135–137 (2001)
29. IEEE Standard 1471-2000: IEEE Recommended Practice for Architectural Description of Software-Intensive Systems. IEEE (2000)
30. Mihaylov, B.: Towards an agile software architecture (2015). <https://www.infoq.com/articles/towards-agile-software-architecture/>. Accessed 31 Aug 2023
31. Viviani, L., Guerra, E., Melegati, J., Wang, X.: An empirical study about the instability and uncertainty of non-functional requirements. In: Stettina, C.J., Garbajosa, J., Kruchten, P. (eds.) *International Conference on Agile Software Development*, vol. 475, pp. 77–93. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-33976-9_6
32. Chen, L., Babar, M.A., Ali, N., Chen, X.: A taxonomy of non-functional requirements for software architecture evaluation. *IEEE Access* **8**, 31753–31769 (2020)
33. Cuesta, C.E., Pérez, J.A., Pérez, M.S.: Dealing with non-functional requirements in model-driven development. In: *Advances in Intelligent Systems and Computing*, pp. 15–27 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Investigating Open Innovation Practices to Support Requirements Management in Software Ecosystems

Paulo Malcher^{1,2}(✉) , Davi Viana³ , Pablo Oliveira Antonino⁴ ,
and Rodrigo Pereira dos Santos¹ 

¹ Federal University of the State of Rio de Janeiro, Rio de Janeiro, Brazil
`malcher@edu.unirio.br`, `rps@uniriotec.br`

² Federal Rural University of Amazônia, Capitão Poço, Brazil

³ Federal University of Maranhão, São Luís, Brazil
`davi.viana@ufma.br`

⁴ Fraunhofer Institute for Experimental Software Engineering, Kaiserslautern,
Germany
`pablo.antonino@iese.fraunhofer.de`

Abstract. Software ecosystems (SECO) affect requirements management when considering multiple actors (i.e., keystone, third-party developer, users) from different organizations using several communication channels such as issue trackers and forums. To deal with this scenario, professionals involved in requirements management in SECO have resorted to several open innovation (OI) practices. Our study aims to investigate OI practices applied to support requirements management in SECO. We conducted a field study based on interviews with 21 professionals involved in requirements management activities in SECO. We identified 10 OI practices to support requirements management in SECO and 14 communication channels to receive/provide requirements from/to external actors. OI practices identified in this study can help practitioners manage requirements in the SECO context in which they are engaged, making this process more informal, open, and collaborative.

Keywords: Open innovation · Requirements management · Software ecosystems · Field study

1 Introduction

Requirements management is a process that captures, traces, manages, and communicates stakeholder needs and changes throughout a project's lifecycle. This process is recognized as fundamental to ensure the delivery of adequate and quality software products [44]. However, new trends in software development, such as software ecosystems (SECO), have presented challenges for requirements management [20]. In SECO, multiple products are derived from a common technological platform based on a central architecture integrating other systems and forming a network of actors and artifacts [26].

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 35–50, 2024.

https://doi.org/10.1007/978-3-031-53227-6_3

The complexity and changing nature of SECO result in several new requirements based on ecosystem trends called emergent requirements that make requirements management difficult [20]. One reason is that multiple actors from different organizations communicate through multiple open communication channels [20]. In this challenging context, professionals involved in requirements management activities in SECO have resorted to open innovation (OI) practices such as co-creation, collaboration, and crowdsourcing.

Several works have addressed the relationship between OI and SECO and requirements engineering (RE) [9, 21, 24, 25]. However, none identified which OI practices have been used to support requirements management in SECO. Implementing external requirements helps continuously to create more value for products and services in SECO [9]. In this work, we aim to investigate the use of OI practices to support requirements management in SECO. To achieve this goal, we conducted a field study based on interviews with 21 professionals involved in related activities in SECO.

Our results show that professionals commonly receive/provide requirements or requirements changes from/to external actors (e.g., customers, users, partners, third-party developers). We also identified that they use 14 communication channels to receive/provide these requirements and 10 OI practices to support requirements management in SECO.

The remainder of this paper is organized as follows: Sect. 2 presents the background and related work; Sect. 3 describes the research method; Sect. 4 presents our results; Sect. 5 present the discussion, implications, and threats to validity; and Sect. 6 concludes the paper with some final remarks.

2 Background and Related Work

Requirements management comprises comprehensive activities that record and maintain evolving requirements [16]. However, it is considered a challenge in SECO [42]. Opening requirements management to external actors is challenging because ecosystem professionals must keep requirements transparent between the keystone and external actors [17]. Hence, SECO represented a radical software engineering (SE) shift, influencing fundamental aspects such as openness, collaboration, and innovation [15, 17]. Linåker and Wnuk [24] state that the OI paradigm may further explain this new context.

OI assumes that companies should use internal and external ideas and paths to market as they look to advance their technology [5]. Moreover, a majority of the innovation within a software has been increasingly reliant on OI [46]. In this scenario, RE needs to take the changes implied in the OI in regard and adapt to them [25]. Several OI practices have been used in software development, such as co-creation, collaboration, and crowdsourcing. These practices are classified into the main OI processes (inbound, outbound, and coupled) [4, 31, 38].

Linåker and Wnuk [24] propose a model for analyzing and managing requirements designed in the context of SECO that clarifies how requirements management can be adjusted to benefit from OI. Fernandez et al. [9] gauged how

common OI is in the RE practice and to what extent it is implemented. For the authors, receiving/providing requirements from/to external actors is common, but implementing requirements in an OI context can be challenging. Linåker et al. [23] propose a model that provides an operational OI perspective on what firms involved in open source SECO (OSSECO) should share, helping them motivate contributions by creating contribution strategies. Our study considers the OI practices cited in the related work presented in this section. Moreover, our study differs from them by investigating the perceptions of professionals involved in requirements management activities in SECO on using the OI practices.

3 Research Method

We conducted a field study as a research method to investigate the use of OI practices to support requirements management in SECO. A field study seeks to investigate how practitioners of some activity deal with the practice or solve problems within their respective contexts [34]. A set of techniques for data collection can be used in a field study, including interviews [33]. Hence, we performed semi-structured interviews based on recommendations for field studies [34] with professionals involved in requirements management activities in SECO.

Our research question (RQ) aimed to allow a researcher to obtain detailed information about participants' experiences, opinions, and perspectives on how they receive/provide requirements or requirements change from/to external actors in SECO and how they manage these requirements in OI context. Our RQ was: *How do OI practices influence requirements management in SECO?*

Data from semi-structured interviews are generally analyzed using qualitative analysis methods [32,34]. We applied coding procedures inspired by the initial Grounded Theory procedures [37] to analyze qualitative data and descriptive statistics to analyze quantitative data. We present the process for conducting the semi-structured interviews and our approach to analyze the results below.

3.1 Semi-structured Interviews

We initially developed an interview guide¹ with interview planning. Afterward, we conducted a pilot interview with one professional involved in requirements management activities in SECO. The pilot checked the questions' clarity and understanding and the estimated time to complete the interview. The pilot participant encouraged us to add the definition of each OI practice presented to clarify possible doubts of the interviewees. We point out that we do not use pilot data in our analysis.

We conducted 21 interviews between July and August 2023 with professionals involved in requirements management activities in SECO. Each interview lasted between 35 and 55 min. We used Google Meet² to record the interviews and

¹ <https://doi.org/10.5281/zenodo.10038855>.

² <https://meet.google.com/>.

Google Docs³ to transcribe them. We transcribed the interviews iteratively, and the researcher coded the interviews, always watching the original video during the coding process even though we automatically transcribed each recording. Hence, we ensured the best and most accurate interpretation possible of each interview. We also fixed errors in the transcripts generated automatically during the coding process. We divided the interviews into three parts:

- i **Characterization of the participants:** We collected information about academic background and experience in industry;
- ii **Presentation of the concepts used in the interview:** We presented the definitions of SECO, requirements management, and OI to ensure clarity and avoid any confusion or ambiguity about the meaning of each term;
- iii **Questions about OI practices to support requirements management in SECO:** We asked participants about their familiarity with OI, whether they receive/provide requirements or requirements change from/to external actors, and how this happens. Finally, we asked what OI practices they use to support requirements management in SECO. In this last question, we used the strategy adopted by Greiler et al. [13]. Such strategy consists of initially obtaining answers without presenting any examples of OI practices (unguided impressions) and so obtaining them after presenting a set of OI practices identified in our related work (guided impressions). This set of OI practices encourages deeper discussion as well as encourages participants to consider practices not immediately remembered.

We adopted the concept of “saturation” to establish the number of interviews required in our study. A study reaches saturation when conducting a new set of interviews does not produce new emerging data [8]. According to Guest et al. [14], saturation can usually be obtained with at least 12 interviews. In our study, we interviewed 21 professionals. We reached saturation with 18 interviews, in line with the work of Guest et al. [14]. In each interview, we observed whether participants repeated earlier discussed topics. Interview recordings and transcriptions were continually revisited in an iterative process. As no new codes or insights emerged in three consecutive interviews, we realized our codes and insights were fully saturated and stopped recruiting new participants.

3.2 Characterization of Participants

We used convenience sampling to select participants for our study based on their being nearby and available [1]. However, we looked for diverse participants in terms of experience and contacted professionals involved in requirements management activities in SECO from our network by email and other communication channels (WhatsApp and LinkedIn). We also used snowball sampling, where early participants referred other professionals to participate in the study. In addition, we applied a questionnaire with the consent form and some questions

³ <https://docs.google.com>.

about the characterization of the participants⁴. All participants have experience in requirements management, SECO, and OI. This helps ensure that the selected sample is representative and relevant to the research goals. We assigned each participant a unique identifier (P1 to P21). Table 1 summarizes the information about the interview participants.

Table 1. Characterization of participants.

ID	Academic background in computer science	Experience in requirements management	Engagement in SECO	Participation in projects that use IO
P1	Master's degree	15 years	Yes	Yes
P2	PhD	25 years	Yes	No
P3	Specialization	10 years	No	No
P4	Specialization	10 years	Yes	Yes
P5	Specialization	3 years	I don't know	I don't know
P6	Master's degree	7 years	Yes	Yes
P7	Specialization	6 years	Yes	Yes
P8	PhD	9 years	Yes	Yes
P9	Specialization	10 years	Yes	No
P10	Master's degree	12 years	Yes	Yes
P11	PhD	5 years	Yes	No
P12	Master's degree	15 years	Yes	Yes
P13	Master's degree	8 years	Yes	No
P14	Master's degree	5 years	Yes	Yes
P15	PhD	30 years	Yes	Yes
P16	Specialization	13 years	Yes	Yes
P17	PhD	2 years	Yes	Yes
P18	PhD	2 years	Yes	Yes
P19	Bachelor's degree	10 years	No	No
P20	PhD	3 years	Yes	No
P21	Master's degree	15 years	Yes	Yes

Six (28,6%) of the 21 participants have between 2 and 5 years of experience in requirements management, eight (38,1%) have between 6 and 10 years, and seven (33,3%) have more than 10 years of experience. Some participants answered “no” or “I don't know” to questions about their engagement in SECO and participation in projects using OI. However, they confirmed involvement in these scenarios when we presented the concepts during the interviews. The participants had been engaged in 11 different SECO. We described⁵ and classified them into proprietary⁶ (7), open source⁷ (3), and hybrid⁸ (1) SECO.

⁴ <https://doi.org/10.5281/zenodo.10038855>.

⁵ <https://doi.org/10.5281/zenodo.10038855>.

⁶ In a proprietary SECO, organizations are concerned with keeping their assets protected by intellectual property [7].

⁷ In an open source SECO, the keystone is an OSS community over a set of projects in an open-common platform [11].

⁸ In a hybrid SECO, open source and proprietary practices are combined [26].

3.3 Coding Process

To analyze the interviews, we initially performed an open coding approach inspired by the initial procedure of the Grounded Theory [37]. During the open coding process, we divided the transcripts into coherent units (sentences or paragraphs) and added **preliminary codes** representing the key points each participant talked about. Subsequently, we defined a set of **focused codes** that captured the most frequent and relevant factors in the participants’ perceptions. After performing open coding, we used axial coding described by Charmaz [3] to group the codes into **categories**. In these steps, we used the Atlas.TI tool⁹ as support to create the codes and categories. Table 2 shows the example of the coding process for one transcript with resulting codes and categories.

Table 2. Illustration of the coding process.

Coherent unit: “We have a collaborative flow in which cooperated members carry out this open innovation. They develop and ship the code to us. We can embed it in our code, but first, we understand, document, and specify that code.” (P11)			
Preliminary code	Focused code	Category	Core category
<i>We have a collaborative flow in which cooperated members</i>	Collaboration	Coupled	OI practice

One researcher conducted and coded the interviews over in iterative cycles. The other three researchers, with more than 15 years in SE, double-checked the results and ensured the compliance of the final dataset. Moreover, we continuously revisited the interview recordings and transcripts in an iterative process.

4 Results

This section presents the results obtained in the semi-structured interviews performed in our field study that investigated the use of OI practices to support requirements management in SECO. We identified that most participants are familiar with OI, although some of them did not know it by such terminology. Moreover, participants use multiple communication channels to receive/provide requirements or requirements change and several OI practices to support activities related to requirements management in SECO. We detail our results next.

4.1 Communication Channels in SECO

We initially asked professionals about their familiarity with the OI concept. This question aimed “to break the ice” and verify the participants’ perceptions about the subject. All participants rated their familiarity with the OI on a scale of 1 to 5, where 1 meant less familiar and 5 meant more familiar. One participant considered himself/herself in level 1, four in level 2, eight in level

⁹ <https://www.atlasti.com>.

3, three in level 4, and six in level 5. We identified many participants were unfamiliar with the term “open innovation”. However, after we explained the concept at the beginning of the interviews, these participants reported they had already participated in projects that used OI. P13 highlighted: “*After your presentation, I realized I am quite familiar with the subject. I did not know it by that name, but I realized that we have this context of innovation in the ecosystem in which I participate*”.

We also asked participants if they usually receive/provide requirements from/to external actors to the projects they have been involved in SECO. If yes, we asked how they received/provided them. In response, 20 of the 21 participants stated that they received/provided requirements or requirements change from/to external actors. Only one participant claimed never to have provided/received requirements or requirements change from/to external actors. However, this participant mentioned during the interview that they use a tool provided by key-stone to clarify doubts, report bugs, interact with SECO members from other organizations, and send suggestions for improvements.

Regarding how the participants receive/provide requirements or requirements change from/to external actors, we identified 14 communication channels. Communication channels are mainly used to improve and maintain a project’s presence in a SECO and ensure that projects share knowledge at the ecosystem level with several contributors distributed geographically that possess different interests [39]. Moreover, communication channels help enhance OI practices, connecting key stakeholders, such as customers, suppliers, or business partners, and collaborating in the development of new products and services [2]. We classify these communication channels into three categories: (i) open online communication channels; (ii) closed online communication channels; and (iii) face-to-face communication channels. We also added the number of participants who cited each code. Table 3 presents the codes and categories resulting from our analysis.

Table 3. Communication channels to receive/provide requirements or requirements change from/to external actors in SECO.

Open online	Closed online	Face-to-face
App store (2)	Emails (7)	Face-to-face meetings (6)
Forums (6)	Feedback systems (5)	Hackathons (1)
Issue/bug trackers (1)	Forms (3)	Product demonstrations (3)
Software repositories (1)	Help desks (3)	Technical visits (1)
	Instant messaging apps (3)	

Open online communication channels facilitate information flows between the multiple actors in SECO [20]. The open communication paradigm in SECO provides opportunities for ‘just-in-time’ RE [19]. Participants cited the use of forums, app stores, issue/bug trackers, and software repositories to

receive/provide requirements or requirements change from/to external actors in SECO. Forums, such as Stack Overflow, were mostly mentioned by the participants. P8 highlighted: “*We are looking at the Stack Overflow and are mapping if there are any requirements around a tool, a product, or a software that we will need to change*”. According to Vevers et al. [43], to fully understand how a SECO works, the community needs to be studied as well, and this can be done by looking at issue/bug trackers and forums.

Closed online communication channels enable fast responses and can speed up decision making [35]. Participants also cited emails, forms, remote meetings, instant messaging apps, feedback systems, and help desks as channels to receive/provide requirements or requirements change from/ to external actors in SECO. Some participants highlighted the use of multiple closed online communication channels. P5 reported: “*For those who were not users of the tool, they contacted us in various ways, official letter, email, and even WhatsApp in an informal way*”. According to Johnson et al. [18], helpful information could be obtained through analysis of these multiple channels in SECO, both by the platform provider and the partner apps in their innovation processes.

Face-to-face communication channels are stimulus rich, i.e., enable the use of senses (auditory, visual, tactile, olfactory, and gustatory) in verbal and nonverbal activities [28]. Participants mentioned face-to-face meetings, product demonstrations at conferences or for other organizations, technical visits, and hackathons to receive/provide requirements or requirements change from/to external actors in SECO. Some participants conducted hackathons to identify requirements from external actors in SECO. P8 shared: “*We run hackathons to obtain requirements that may be important for new products or products already on the organization’s roadmap*”. According to Valença et al. [40], a hackathon can be seen as a strategy to support SECO evolution, enabling a company to gather new developers for its ecosystem, assess the software platform by identifying bugs, and verify to what extent the requirements for applications are fulfilled.

4.2 OI Practices to Support Requirements Management in SECO

Our main objective was to identify OI practices to support requirements management in SECO through interviewing professionals. As described in Sect. 3, we iteratively coded their responses to the question: “*What open innovation practices have you used to support requirements management activities in software ecosystems?*” and grouped them into categories. Thus, we identified ten OI practices that support requirements management in SECO (Table 4). We identified eight OI practices in the unguided impressions, i.e., at least one participant mentioned the OI practice before we presented the set of OI practices. Only two OI practices (open source and cooperation) were mentioned exclusively in the guided impressions.

We categorize OI practices according to OI processes (inbound, outbound, and coupled) [38]. Inbound OI seeks knowledge from external sources (e.g., suppliers, customers, competitors, and partners). Outbound OI explores internal

knowledge externally. Coupled OI is a process where knowledge can flow inbound and outbound through active collaboration with partners to innovate. Table 4 shows the ten OI practices used to support requirements management in SECO, their categories, and the total number of participants that cited them. Below, we detail the OI practices identified in the study.

Table 4. OI practices to support requirements management in SECO.

Outbound	Inbound	Coupled
Open source (6)	Crowdsourcing (7)	Collaboration (17)
Venturing (1)	Hackathon (8)	Co-creation (6)
	University research grants (2)	Coopetition (3)
	Customer immersion (2)	
	Outsourcing R&D (2)	

Customer immersion is a collaborative innovation practice that focuses on the customer’s experience of using products or services [38]. Participants highlighted intense interaction with customers at events or agile ceremonies to identify requirements or requirements change. According to Gassmann [12], customer involvement is the principal constituent of OI. P18 mentioned: *“For more important customers, they sent invitations to events where they would expose the platform or software and received feedback them”*.

Hackathons are events with an element of competition, where participants work in teams over a short period to ideate, collaborate, design, rapidly prototype, test, iterate, and pitch their solutions to a determined challenge [10]. Some participants stated that hackathos are OI practices that support requirements management in SECO. These participants mentioned that they carry out or participate in hackathons to identify ideas, emerge and define requirements, create synergy between partners, and train different SECO actors. Hackathon is one key practice to enable OI [10]. P6 mentioned: *“When I want ideas or to understand a topic, I organize hackathons. Hackathon is cool because we listen to several ideas and select them”*.

Crowdsourcing consists of outsourcing processes, traditionally carried out internally, to an indefinite, generally large group of people [38]. Participants mentioned that crowdsourcing allows several SECO actors to contribute to requirements management. P1 stated: *“We have crowdsourcing when several groups come together. Our ventures come together to fund ideas”*. P18 commented: *“We used crowdsourcing to let the crowd say what was best about the system”*.

Outsourcing R&D consists of R&D services hiring from other organizations [41]. Participants said they worked in organizations that provided R&D services to keystone. P9 highlighted: *“The company I work for was hired as responsible for credit-related systems. When I need to request a change in systems not under our supervision, for example, customer or internal code systems.*

I speak with [omitted] (keystone), but not with the companies responsible for these systems”.

University research grants consist of funding external research projects by researchers and scientists in universities to access external knowledge [4]. Some participants shared that keystone offered research grants for SECO members to carry out requirements management activities. P9 shared: *“The government has a digital transformation project that has injected resources into [omitted] (keystone). So, [omitted] (keystone) opened a call for grants for analysts and developers from the other organizations that are part of the ecosystem to work on the development of some module”.*

Venturing is defined as starting up new organizations drawing on internal knowledge, and possibly also with finance, human capital, and other support services from your enterprise [41]. Some participants reported that the companies they work for sometimes create new companies to meet specific requirements of the common technological platform or customers. P1 claimed: *“We have a group of ventures that support each other for innovation initiatives and initiatives to meet requirements and provide solutions for customers”.*

Open source aims to reveal internal technologies without immediate financial rewards for indirect benefits to the company [45]. Some participants highlighted that they identify changes to product requirements they develop by participating in open source initiatives. P14 reported: *“I participated in a project that used open source last year. We had an algorithm that made this automatic match between investors and startups. So, we helped other developers because it was something nobody could do, and our company got feedback”.*

Co-creation refers to the contribution provided by the consumer to the process of creating value for the company, allowing the consumer to actively contribute to designing, analyzing, controlling, and evaluating products and processes [38]. Some participants commented on the active participation of customers in requirements management in SECO. P1 shared: *“We have some key customers who contribute to our activities and give us feedback”.* P14 stated: *“We are a design-driven company. Co-creation is what we do”.*

Collaboration involves internal resources operating in different business areas and extends to integrating external resources to define and develop innovative projects [38]. Several participants mentioned that collaborating with other organizations allows identifying requirements change, clarifying doubts, and implementing new features. P10 shared: *“A partner institution came to us so that we could clarify some doubts about the functioning of the systems and make some business comparisons to implement new functionalities”.*

Coopetition is characterized by a balance between cooperative and competitive forces [6]. Some participants reported that there are direct and indirect partnerships between competitors in SECO. Thus, some organizations need to compete in requirements prioritization. P18 mentioned: *“I observed coopetition when there were conflicting requirements between keystone’s partners. They were indirect partners because they evolved the platform and used each other’s solutions. However, they competed when it came to developing and sending add-ons”.*

5 Discussion

From the answers obtained in 21 interviews with professionals who carry out requirements management activities in SECO, we identified how these professionals receive/provide requirements or requirements change from/to external actors and which OI practices are used to support these activities. We discuss our main results next.

Regarding the **communication channels** used to receive/provide requirements or requirements change from/to external actors in SECO, we identified that professionals use open online, closed online, and face-to-face communication channels. The relationship between open communication channels, requirements and SECO has already been investigated in the literature [20, 22]. Knauss et al. [20] state that open communication channels allow transparent communication between developers and customers and are important for exploring RE practices in SECO. Linaker et al. [22] mentioned open communication channels, open requirements management, and active ecosystem engagement as resources to enable an open collaboration in SECO. Hence, open communication channels allow OI practices that influence open requirements management in SECO.

In our study, P8 cited that he analyzes forums such as Stack Overflow to identify possible requirements. In the same direction, Knauss et al. [20] stated that some internal stakeholders even actively track open communication channels of other actors to identify crosscutting problems without this task being formally assigned to them. For the authors, open communication channels have shown their value for building communities over healthy ecosystems. Moreover, these channels offer an exciting opportunity to improve scalability by facilitating decentralized “just-in-time” RE and supporting agile development.

Regarding **OI practices** to support requirements management in SECO, we observed that SECO and OI are related mainly to collaboration between different actors (including external actors) over a common technological platform. Jansen [17] defines OI as a focus area of SECO governance. The OI focus area is concerned with sharing knowledge across the ecosystem to feed external developers with new possibilities for improvement, also known as niche creation [17]. Hence, the OI focus area directly relates to requirements management.

Our results also show that OI practices influence how requirements management is carried out. Fernandez and Svensson [9] stated that OI as part of the RE process is becoming more and more fully explored from both the inbound and outbound. Several participants of our study highlighted the informality of OI practices to support requirements management in SECO. Linaker and Wnuk [24] considered RE in OI and presented the open RE concept. The open RE is informal, transparent, decentralized, distributed, and collaborative [24]. According to the authors, open RE is informal to different degrees, including the level at which requirements are managed.

5.1 Implications for Practitioners and Researchers

Implications for Practitioners. First, practitioners can identify in this study communication channels used to receive/provide requirements or requirements change from/to external actors in SECO. This can assist them in the development of strategies for using these communication channels to identify requirements or requirements change in the SECO they participate. Second, practitioners can identify in this study OI practices used to support requirements management in SECO. Hence, they can analyze whether they can use them in their context.

Implications for Researchers. We also identified implications for researchers in our study. First, the set of communication channels used to receive/provide requirements or requirements change from/to external actors in SECO identified in this study can be useful to researchers investigating requirements flows in SECO. Second, the set of OI practices to support requirements management in SECO presented in this study can be investigated in the context of other RE activities in SECO. Moreover, it can also be useful in research on emergent RE contexts such as crowd-based RE, open RE, and cross-domain RE.

5.2 Threats to Credibility and Reliability

In contrast to quantitative studies, qualitative studies are more prone to threats to credibility than to validity [13,29]. The matters of validity and reliability in qualitative research rely on the meticulousness, thoroughness, and honesty employed by the researchers throughout the data collection and analysis processes [30]. Thus, we outline the potential threats to external and internal credibility in the following.

Internal credibility refers to the credibility of interpretations and conclusions within the underlying setting or group [27]. Interpretive validity is a potential threat to the internal credibility of this study. During interviews and transcripts, there is a risk that researchers will impose their interpretations rather than understand participants' perspectives. We mitigated this threat by asking clear questions to participants and encouraging them to reflect deeply on their answers. In addition, while the first author of this study did the main coding, the other three authors, with more than 15 years in SE, were extensively involved in cross-checking the results and ensuring the compliance of the final dataset.

External credibility refers to the degree to which the findings of a study can be generalized across different contexts [27]. The number and experience of interviewed participants are a potential external threat to this study. We mitigate this following the same strategy of other works [13,29,36] that conducted field studies with software developers. These works considered the recommendations of Guest et al. [14] that saturation in semi-structured interviews can be achieved with at least 12 interviews. Hence, we conducted interviews until we reached saturation. We conducted 21 interviews, and we emphasize that no new categories or codes emerged in the last three interviews, indicating that saturation

was reached. In addition, we selected professionals with different background and experience in requirements management activities in SECO. This contributed to a more significant variety of information with different perspectives.

6 Conclusion and Future Work

This paper addressed the following RQ: “*How do OI practices influence requirements management in SECO?*”. We performed a field study based on interviews with 21 professionals to investigate the OI practices used to support requirements management in SECO. We identified that the use of multiple open communication channels by internal and external actors allows different OI practices, such as hackathons, crowdsourcing, co-creation, collaboration, and open source, which provides knowledge sharing across the ecosystem. Hence, we conclude that OI practices affect requirements management in SECO, making it more informal, open, and collaborative. As future work, we can investigate the impact of specific OI practices on requirements management in SECO, such as crowdsourcing. We plan to identify how crowd feedback affects requirements management in SECO. Furthermore, future work should consider the impact of the different types of SECO (open, proprietary, or hybrid) for using OI practices to support requirements management.

References

1. Baltes, S., Ralph, P.: Sampling in software engineering research: a critical review and guidelines. *Empir. Softw. Eng.* **27**(4), 94 (2022)
2. Cepeda-Carrion, I., Ortega-Gutierrez, J., Garrido-Moreno, A., Cegarra-Navarro, J.G.: The mediating role of knowledge creation processes in the relationship between social media and open innovation. *J. Knowl. Econ.* **14**(2), 1275–1297 (2023)
3. Charmaz, K.: *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage Publications (2006)
4. Chesbrough, H., Brunswicker, S.: *Managing open innovation in large firms*. Technical report. Garwood Center for Corporate Innovation at California University, Berkeley in US & Fraunhofer Society in Germany (2013)
5. Chesbrough, H.W.: *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business Press (2003)
6. Corbo, L., Kraus, S., Vlačić, B., Dabić, M., Caputo, A., Pellegrini, M.: Coopetition and innovation: a review and research agenda. *Technovation* **122**, 102624 (2023)
7. Costa, L.A., Fontão, A., Santos, R.: Toward proprietary software ecosystem governance strategies based on health metrics. *IEEE Trans. Eng. Manage.* **69**(6), 3589–3603 (2022)
8. Creswell, J.W., Poth, C.N.: *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage Publications (2016)
9. Fernandez, S., Svensson, R.B.: A survey of practitioners use of open innovation. In: *43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pp. 305–312. IEEE, Vienna, Austria (2017)

10. Flores, M., et al.: How can hackathons accelerate corporate innovation? In: Moon, I., Lee, G.M., Park, J., Kiritsis, D., von Cieminski, G. (eds.) APMS 2018. IAICT, vol. 535, pp. 167–175. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99704-9_21
11. Franco-Bedoya, O., Ameller, D., Costal, D., Franch, X.: Open source software ecosystems: a systematic mapping. *Inf. Softw. Technol.* **91**, 160–185 (2017)
12. Gassmann, O.: Opening up the innovation process: towards an agenda. *R&D Manag.* **36**(3), 223–228 (2006)
13. Greiler, M., Storey, M.A., Noda, A.: An actionable framework for understanding and improving developer experience. *IEEE Trans. Software Eng.* **49**(4), 1411–1425 (2022)
14. Guest, G., Bunce, A., Johnson, L.: How many interviews are enough? An experiment with data saturation and variability. *Field Methods* **18**(1), 59–82 (2006)
15. Hanssen, G.K., Dybå, T.: Theoretical foundations of software ecosystems. In: IWSECO@ ICSOB, pp. 6–17 (2012)
16. ISO/IEC/IEEE29148: ISO/IEC/IEEE international standard - systems and software engineering - life cycle processes - requirements engineering. ISO/IEC/IEEE 29148:2018(E), pp. 1–104 (2018)
17. Jansen, S.: A focus area maturity model for software ecosystem governance. *Inf. Softw. Technol.* **118**, 106219 (2020)
18. Johnson, D., Tizard, J., Damian, D., Blincoe, K., Clear, T.: Open CrowdRE challenges in software ecosystems. In: 4th International Workshop on Crowd-Based Requirements Engineering (CrowdRE), pp. 1–4. IEEE, Zurich, Switzerland (2020)
19. Knauss, E., Damian, D., Knauss, A., Borici, A.: Openness and requirements: opportunities and tradeoffs in software ecosystems. In: 22nd International Requirements Engineering Conference, pp. 213–222. IEEE, Karlskrona, Sweden (2014)
20. Knauss, E., Yussuf, A., Blincoe, K., Damian, D., Knauss, A.: Continuous clarification and emergent requirements flows in open-commercial software ecosystems. *Requirements Eng.* **23**(1), 97–117 (2018)
21. Linåker, J., Munir, H., Runeson, P., Regnell, B., Schrewelius, C.: A survey on the perception of innovation in a large product-focused software organization. In: Fernandes, J.M., Machado, R.J., Wnuk, K. (eds.) ICSOB 2015. LNBIP, vol. 210, pp. 66–80. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19593-3_6
22. Linåker, J., Runeson, P.: How to enable collaboration in open government data ecosystems: a public platform provider’s perspective. *JeDEM-eJournal eDemocracy Open Gov.* **13**(1), 1–30 (2021)
23. Linåker, J., Munir, H., Wnuk, K., Mols, C.: Motivating the contributions: an open innovation perspective on what to share as open source software. *J. Syst. Softw.* **135**, 17–36 (2018)
24. Linåker, J., Wnuk, K.: Requirements analysis and management for benefiting openness. In: IEEE 24th International Requirements Engineering Conference Workshops (REW), pp. 344–349. IEEE, Beijing, China (2016)
25. Linåker, J., Regnell, B., Munir, H.: Requirements engineering in open innovation: a research agenda. In: 2015 International Conference on Software and System Process, pp. 208–212. ACM, New York, NY, USA (2015)
26. Manikas, K.: Revisiting software ecosystems research: a longitudinal literature study. *J. Syst. Softw.* **117**, 84–103 (2016)
27. Onwuegbuzie, A., Leech, N.: Validity and qualitative research: an oxymoron? *Qual. Quant.* **41**, 233–249 (2007)

28. Raybourn, E.: Strategies for working remotely: responding to pandemic-driven change with cross-organizational community dialog. In: Fui-Hoon Nah, F., Siau, K. (eds.) *HCI in Business. Government and Organizations*, vol. 13327, pp. 407–416. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-05544-7_30
29. Ribeiro, B.B., Costa, C., dos Santos, R.P.: Understanding and analyzing factors that affect merge conflicts from the perspective of software developers. *J. Softw. Eng. Res. Dev.* **10**, 12:1–12:17 (2022)
30. Robson, C.: *Real World Research: A Resource for Social Scientists and Practitioner-Researchers* (2002)
31. Rumanti, A.A., Rizana, A.F., Ramadhan, F., Reynaldo, R.: The impact of open innovation preparation on organizational performance: a systematic literature review. *IEEE Access* **9**, 126952–126966 (2021)
32. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empir. Softw. Eng.* **14**, 131–164 (2009)
33. Seaman, C.: Qualitative methods in empirical studies of software engineering. *IEEE Trans. Software Eng.* **25**(4), 557–572 (1999)
34. Singer, J., Sim, S.E., Lethbridge, T.C.: Software engineering data collection for field studies. In: Shull, F., Singer, J., Sjøberg, D.I.K. (eds.) *Guide to Advanced Empirical Software Engineering*, pp. 9–34. Springer, London (2008). https://doi.org/10.1007/978-1-84800-044-5_1
35. Souza, A., Silva, S.T., França, J., Dias, A., Oliveira, J., Vivacqua, A.: Communication channels and their challenges: an analysis of software development teams during the covid-19 pandemic. *ACM Trans. Comput.-Hum. Interact.* **7**(GROUP) (2022)
36. Steglich, C., et al.: Factors that affect developers’ decision to participate in a mobile software ecosystem. *J. Syst. Softw.*, 111808 (2023)
37. Strauss, A., Corbin, J.: *Basics of Qualitative Research*. Sage Publications (1990)
38. Strazzullo, S., Cricelli, L., Grimaldi, M., Ferruzzi, G.: Connecting the path between open innovation and industry 4.0: a review of the literature. *IEEE Trans. Eng. Manag.*, 1–13 (2022)
39. Tantisuwankul, J., et al.: A topological analysis of communication channels for knowledge sharing in contemporary GitHub projects. *J. Syst. Softw.* **158**, 110416 (2019)
40. Valença, G., Lacerda, N., Rebelo, M.E., Alves, C., de Souza, C.R.B.: On the benefits of corporate hackathons for software ecosystems – a systematic mapping study. In: Franch, X., Männistö, T., Martínez-Fernández, S. (eds.) *PROFES 2019. LNCS*, vol. 11915, pp. 367–382. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35333-9_27
41. van de Vrande, V., de Jong, J.P., Vanhaverbeke, W., de Rochemont, M.: Open innovation in SMEs: trends, motives and management challenges. *Technovation* **29**(6), 423–437 (2009)
42. Vegendla, A., Duc, A.N., Gao, S., Sindre, G.: A systematic mapping study on requirements engineering in software ecosystems. *J. Inf. Technol. Res. (JITR)* **11**(1), 49–69 (2018)
43. Ververs, E., van Bommel, R., Jansen, S.: Influences on developer participation in the Debian software ecosystem. In: *International Conference on Management of Emergent Digital EcoSystems*, pp. 89–93. ACM, New York, NY, USA (2011)
44. Wibowo, A., Davis, J.: Requirements traceability ontology to support requirements management. In: *Australasian Computer Science Week Multiconference*. ACM, New York, NY, USA (2020)

45. Wnuk, K., Pfahl, D., Callele, D., Karlsson, E.A.: How can open source software development help requirements management gain the potential of open innovation: an exploratory study. In: ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, pp. 271–280. ACM, New York, NY, USA (2012)
46. Wnuk, K., Runeson, P.: Engineering open innovation – towards a framework for fostering open innovation. In: Herzwurm, G., Margaria, T. (eds.) ICSOB 2013. LNBP, vol. 150, pp. 48–59. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39336-5_6

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Requirements Tool Practices that Drive Business Agility

Andreas Birk^(✉)

Software.Process.Management, Usedomstr. 15, 70439 Stuttgart, Germany
andreas.birk@swpm.de

Abstract. CONTEXT: Successful agile teams advance their work practices continuously. The continuous improvement of effective tool-based requirements practices is an important foundation of business agility. However, requirements tool practices are still widely rooted in plan-based approaches. They are not yet suited well for agile teams or agile businesses. OBJECTIVE: Report and make available an approach for continuous improvement of requirements practices so that tool-based requirements management can drive business agility. METHOD: Industry experience report based on a series of cases from different sources, including ones with involvement of the author. RESULTS: Processes and work practices for evolutionarily introducing and adapting requirements tools and tool-based requirements practices, in a way that supports business agility. CONCLUSION: The presented practices can guide organizations towards establishing effective, tool-based requirements practices that support business agility. A foundation is laid for further systematic investigation and development of the approach.

Keywords: Product Management · Software Requirements · Requirements Tools · Business Agility

1 Business Agility, Requirements, and Tools

Business agility is a very intuitive concept that guides the vision of modern product management and product development. While a single authoritative definition is lacking, the concept is generally associated with *the ability to rapidly and systematically adapt to market, environmental, and technological changes* (cf. [1]).

Business agility can be viewed as an extension of Lean Startup [2] into established, non-startup business environments: Like in Lean Startup, agile development is the driving force for finding and maintaining viable business models (cf. [3]). Agile development is guided by the ideas of customer value and business value (cf. [4, 5]).

Requirements practices are an important foundation of business agility. Product management uses requirements for capturing market changes and customer demands, and for communicating these to development. Development transforms requirements into new product versions that shall create future business value (see Fig. 1).

Mature development organizations, regardless of whether plan-based or agile, have the capability to continuously improve their management and development practices.

For instance, in Scrum the role of the Scrum Master (core task: remove impediments) and the practices Daily Scrum meeting, and Sprint Retrospective serve the purpose of continuous improvement [4].

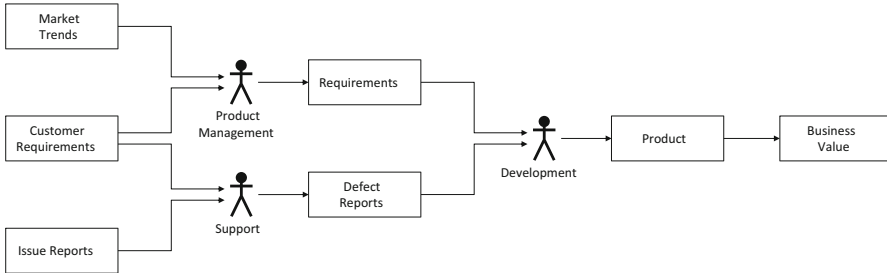


Fig. 1. The role of requirements in mediating between customer need and business value.

Continuous improvement is also key to maintaining business agility and its requirements practices. However, because requirements practices in larger and more complex environments must be tool-based, specific challenges emerge: Since the traditional generations of requirements tools are firmly rooted in plan-based development approaches, there is little guidance and support for the continuous or even agile evolution and improvement of tool-based requirements practices.

This paper wants to contribute to overcoming this limitation. It is a long-term industrial experience report that proposes a new process for evolutionarily improving tool-based requirements practices in a way that supports business agility. The process can be applied for further optimizing work practices on an existing tool platform as well as for introducing a new requirements tool and suitable associated work practices.

The next sections introduce key characteristics of requirements tools, the requirements tools market, and the state of tool-based requirements practices (Sect. 2), propose the process for continuously evolving tool-based requirements practices (Sect. 3) and provide an experience-based justification (Sect. 4). The final section points out conclusions, discusses empirical evidence, and proposes future work (Sect. 5).

2 Requirements Tools

Requirements practices usually require suitable tool support to be effective and efficient. Modern requirements practices therefore encompass the processes and their associated tool support. Both must be considered as a unit (i.e., *tool-based requirements practices*). The tools most widely used are desktop office applications, in particular text processors and spreadsheet tables. They have severe limitations, namely limited central availability (single point of truth) and little support for versioning and tracing.

Specialized requirements tools are available since the 1990s, first as client/server solutions, later as web applications and now increasingly often as cloud applications (SaaS, Software as a Service). Initially, they supported specification-based requirements management. Most modern tools also support agile requirements workflows. DeGea

et al., Wiegers, and Bühne and Herrmann (IREB) provide overviews of the tool market and tool functionality [6–8].

The characteristics of the initial client/server tool generation (e.g., huge expensive products, difficult to install and access) still dominate and bias our today’s perception of requirements tools and how we deal with them. This is particularly true for the selection and introduction of requirements tools and tool-based requirements practices.

Figure 2 shows a typical tool selection and introduction process as it can be found across industry and in the literature (cf. [6–8]). It is built on the comparative evaluation of several candidate tools, in a two-step process (longlist and shortlist evaluation), usually involving checklist scoring, vendor demos, open trial-uses, and vendor-driven proofs of concept (PoC). The selection processes can last many months, sometimes up to a few years. The associated requirements work practices are often based on the vendors’ blueprints, with the later users involved very little into process design. As consequences, effective tool-based requirements practices are still rare. Their contributions to business agility fall far short of what would be possible.

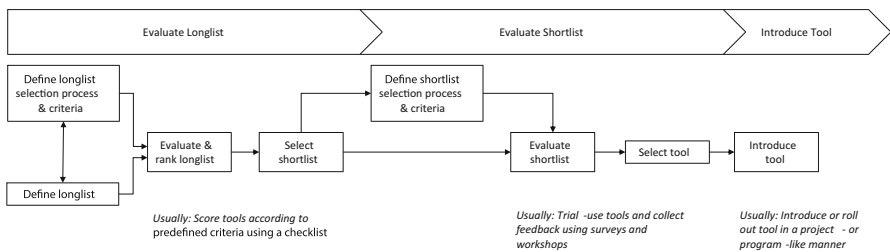


Fig. 2. A typical traditional requirements tool selection process.

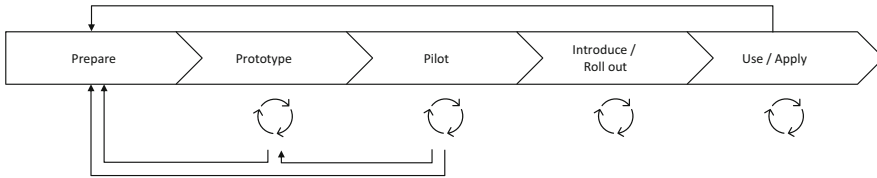
Today’s tool generation allows for new requirements work practices and for more efficient evolutionary improvement approaches: Cloud applications can be accessed very easily for trial-usage. Powerful administration functionality and cloud and virtualization technology allow for easily switching between different candidate solutions. These developments enable new ways for designing and deploying tool-based requirements practices. The following section proposes such a process.

3 Continuously Evolving Tool-Based Requirements Practices

A process for evolving tool-based requirements practices must be iterative, staged, focused, and collaborative: Sufficiently small iterations foster rapid progress and alignment, reduce risk of failure, and fit with Agile. Stages allow for controlled addition of complexity. Focus through objectives and scope gives success criteria and alignment. Collaboration reduces overhead, supports alignment, and, again, fits in well with Agile.

The proposed process has five steps: Prepare, Prototype, Pilot, Introduce/Roll out, and Use/Apply (see Fig. 3). Table 1 describes the activities of each step, their results, and the key actors involved. The actors are:

Core Team: The persons running the improvement project. The group should be small and include all relevant perspectives, usually: Requirements experts (i.e., methods, processes), tool experts (i.e., how to best support practices by the given tool), and stakeholder experts familiar with the application contexts of the practices and tools (e.g., product managers, business analysts, or IT operations managers).



The two phases *Pilot* and *Introduce / Roll out* may be merged into one, especially with smaller solution elements or in smaller contexts. Switch or fundamentally change, update, or extend the solution if necessary, by moving back to the *Prepare* phase.

Fig. 3. Evolutionary improvement of tool-based requirements practices: Process overview.

Sponsor: The persons who have a key interest that the improved solution becomes available, and who provide the needed budget and organizational support.

Key Stakeholders: A focus group of persons from the target group that shall later apply the improved tool-based requirements practice, and who actively support the development of the solution.

Pilot Stakeholders: A focus group of persons from the later application stakeholders who are willing in trial-using the new solution. Pilot stakeholders should not be key stakeholders in order to be unbiased.

Figure 3 shows the main feedback and iteration relations. Usually, the steps are conducted in sequence, with as many small internal iteration cycles as needed. Feedback occurs mainly from the *Prototype* and *Pilot* steps, if the solution turns out not sufficiently mature or ineffective. Then even the entire project may be stopped.

Once the pilot has been successful, the solution will eventually be made available for common application. Additional adjustments can mostly be made without larger intervention. In larger endeavors, like the introduction of a new tool platform, the entire process may be iterated multiple times with increasing scopes.

Table 1. Evolutionary improvement of tool-based requirements practices: process details.

Process Step	Activities	Main Results	Key Actors
Prepare	Define objectives and scope; identify overall approach of solution development; analyze need and solution options, research, experiment with, and evaluate solution options	Objectives, scope, overall approach	Core Team, Sponsor, Key Stakeholders
Prototype	Design and collaboratively trial use a solution; decide about pilot application	Proven solution in sandbox environment	Core Team, Key Stakeholders
Pilot	Make available the solution to one or more pilot application contexts; guide and support pilot application, adjust the solution if necessary; decide about introduction / roll out	Proven solution in (close to) real environment	Pilot Stakeholders, Core Team
Introduce / Roll out	Make available the solution for common application; train, guide, and support the stakeholders, adjust the solution if necessary	Solution available and applied in real environment	Core Team
Use / Apply	Support the stakeholders on a regular basis; adjust the solution if necessary	Solution generally available	Core Team

4 Experiences and Justification of the Approach

The process has been developed gradually over many projects in various organizations. It shall make this experience available for future improvement projects. Also many additional observations and experience reports from third parties were included. A detailed systematic substantiation of the process cannot be given in this short experience report. However, the following two example cases illustrate how the process was derived and justified, and how it can be conducted in practice.

The first example is a smaller-scale improvement of a tool-based requirements practice. It took place within the established tool-based requirements workflow for the development of large software-controlled high-tech machinery. Sometimes engineers tended to overlook requirements status updates (e.g., from *Defined* to *Approved for Implementation*). It was decided that requirements status transitions should be marked in the tool-internal comments thread of each requirement. Initial research (*Prepare* phase) showed that a ready-to-use solution did not exist (e.g., neither a tool configuration option nor a third-party plug-in). However, a custom workflow script could be implemented easily. It was developed in a sandbox project and tested successfully (*Prototype* phase). Pilot application happened in the productive tool environment under special supervision by the core team. The change was soon released for general use. The entire improvement project was conducted within two weeks.

The second example is a large-scale substitution of an established tool-based requirements process by a new tool from the latest tool generation and with advanced work practices. It happened at a large global product division in the semiconductor industry, with several hundred development staff, over a period of about 1.5 years. The core team included persons from the established requirements management team and the requirements tool's product owner from IT operations.

Each step from the improvement process above could be identified, involving several sub-steps and taking several months. For instance, the *Prepare* step included a systematic study of future tool performance. *Prototyping* involved the design of new tool-based practices across various workshops with key stakeholders like marketing, requirements, and architecture. *Pilot* projects tested the new practices and tried the highly critical migration approach. *Roll out* included comprehensive training activities.

The entire improvement project progressed in a well-controlled manner. The new tool and the new tool-based requirements practices received high acceptance.

5 Conclusions, Evidence, and Future Work

The main conclusions from developing and using the proposed process are: Tool-based requirements practices can be evolved and improved continuously in ways that align well with the iterations and improvement practices of agile methods. Product organizations can strengthen their capabilities to react to market trends and customer demands by continuously advancing tool-based requirements practices. This potentially increases the business value of the organization's products and fosters its business agility.

The process has been presented here as a long-term industrial experience report. Basic substantiation has been provided by two example application cases. Many similar projects influenced the design of the process since the early 2000s until mid 2023. They took place in a wide variety of contexts: Product organizations and internal IT, from small teams to divisions of large corporations, hardware/software products as well as marketed software applications. The author of this experience report was mostly involved in the role of a consultant (i.e., a typical position to provide tool-based guidance and support). So, method development has been performed as a kind of action research. Author bias may have been mitigated, because projects were conducted in teams, and various stakeholders strongly influenced the projects' processes. Experience reports from other sources were

considered, too. For instance, the incremental, staged approach by Rathod, Cebulla and Kugele [9] using which they developed advanced requirements traceability support can be mapped fully on the proposed process.

Future work shall be conducted for systematically substantiating the proposed process. It should also investigate in more detail how the evolutionary improvement of tool-based requirements practices advances agile development effectiveness and business agility. Derived experiences shall be integrated into future versions of the process, in order to provide additional and more detailed methodological support and guidance.

References

1. Business Agility. <https://scaledagileframework.com/business-agility>
2. Ries, E.: The Lean Startup: How Constant Innovation Creates Radically Successful Businesses. Portfolio Penguin, London (2011)
3. Blank, S.: Why the lean start-up changes everything (2013). <https://hbr.org/archive-toc/BR1305>
4. Sutherland, J., Schwaber, K.: The Scrum Guide™: The Definitive Guide to Scrum: The Rules of the Game. Scrum.org and ScumInc. (2020)
5. Scaled Agile, Inc.: Scaled Agile Framework® (SAFe®). <https://scaledagileframework.com>
6. de Gea, J.M.C., Ebert, C., Hosni, M., Vizcaino, A., Nicolas, J., Fernandez-Aleman, J.L.: Requirements engineering tools: an evaluation. *IEEE Softw.* **38**, 17–24 (2021). <https://doi.org/10.1109/MS.2021.3058394>
7. Wiegers, K., Beatty, J.: Software Requirements. Microsoft Press, Redmond, WA (2013)
8. Bühne, S., Herrmann, A.: Handbook Requirements Management. IREB International Requirements Engineering Board e.V. (2022)
9. Rathod, V., Cebulla, T., Kugele, S.: Traceability evaluation in requirements engineering according to automotive SPICE. Presented at the 2023 IEEE 31st International Requirements Engineering Conference (RE) (2023)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Software Procurement



On Public Procurement of ICT Systems: Stakeholder Views and Emerging Tensions

Reetta Ghezzi^(✉) and Tommi Mikkonen

University of Jyväskylä, Jyväskylä, Finland
{reetta.k.ghezzi,tommi.j.mikkonen}@jyu.fi

Abstract. The public sector is a significant consumer of ICT systems. In countries like Finland, where openness, objectivity, and fairness in public acquisitions are deemed essential, public ICT procurement is based on tenders initiated by public sector organizations. The tendering process is regulated by laws that aim to eliminate unfair advantages and provide all potential stakeholders with similar opportunities to participate. However, depending on the stakeholders' perspectives, they may interpret the tendering process differently, leading to tensions among them. In this paper, we examine Finland's public procurement of ICT systems using semi-structured interviews as our data collection method and analyze the results thematically. The interviewees include individuals familiar with tendering and acquisition processes in public organizations and those involved in delivering systems as vendors, representing two different perspectives on the tendering process. The results indicate that although there are significant differences in maturity among public sector organizations participating in procurement, several common themes emerged from nearly all the interviews. Furthermore, in light of contrasting views between public organizations and vendors, recurring tensions arise due to different interpretations of acquisition laws.

Keywords: Public Procurement · Public Sector Software · ICT Procurement · Software Acquisition

1 Introduction

Increasingly, the digital society has led to a growing demand for a wide range of public digital services. For example, Finland has initiated a program with the goal of creating Digital Twins for citizens to improve the targeting of services precisely when they are needed most [11]. This signifies that society is becoming progressively more reliant on Information and Communication Technology (ICT) in general, and software in particular.

The public sector acquires software for public use, a process mandated by EU and national procurement legislation within the EU [1]. The EU and national legislations governing this procurement process aim to ensure equality, transparency, and the consideration of both price and quality with relative weights

[17]. In this context, a public organization initiates the procurement process by issuing a call for tenders. During this tendering process, information system providers compete to offer software solutions that best meet the specified requirements.

Despite procurement laws and national standards, much human judgment plays a role. Consequently, it is not uncommon for disputes related to public procurement, including differing perspectives on the tendering process, specifications, and deals, to end up in court.

In this paper, we investigate stakeholder perspectives regarding the public procurement of ICT systems in Finland. We employed semi-structured interviews, targeting individuals with knowledge of the tendering and acquisition processes within public organizations. We conducted a total of 12 interviews, involving representatives from five public organizations and four vendors engaged in ICT procurement. While some stakeholders share certain projects, not all stakeholders are involved in every project. This work extends a previous Master's thesis on Economics [5], which explored various aspects of public procurement in Finland. In this paper, we focus on stakeholder viewpoints and the tensions that arise from them, with the technical findings falling outside the scope of this study.

The rest of this paper is structured as follows. In Sect. 2, we provide the necessary background for the paper. In Sect. 3, we introduce the applied research approach. In Sect. 4, we present the results of the work, and in Sect. 5, we provide an extended discussion of the results, together with some remarks on the study's limitations. Finally, in Sect. 6, we draw some final conclusions.

2 Background and Motivation

Public agencies that acquire information systems typically expect the system to serve the agency without significant changes for an extended period [12]. This long-term stability often leads to collaborative relationships between public agencies and ICT vendors. Various forms of collaboration exist (e.g., [6]), and public procurement processes define how software systems are acquired. These regulated public procurement procedures aim for non-discriminatory treatment of vendors [8]. However, ICT procurement projects frequently exceed their original schedules and budgets, and planned systems may even be abandoned before project completion [19].

Tendering is the process where an agency in need of a software system solicits bids for projects with fixed or nearly fixed deadlines [12]. The process commences with a description of the problem the acquiring agency faces and the creation of a project proposal to address the issue, often in the form of a Request for Information (RFI). An RFI is a formal method for collecting information from potential suppliers of goods or services. Following an RFI, the next step is the Request for Proposal (RFP), which asks vendors to propose solutions to the customer's problems or business requirements. An RFP is a comprehensive, detailed document that contains all the necessary information for an informed purchasing

decision. Finally, a Request for Quotation (RFQ) can be used to invite suppliers or contractors to submit price bids for standardized products or services produced in repetitive quantities.

Like any software specification, tendering-especially RFP, but to some extent, RFI and RFQ-forms the fundamental description of the resulting IT project. However, public ICT procurement is often challenging due to the specific parameters set for public procurement [16]. Strict control practices and the current methods of procurement units can hinder innovation and cost-effectiveness in public procurement [3].

In particular, it has been argued that EU and national regulations in Finland can impede the effective procurement process [10]. However, strict parameters in public procurement exist for valid reasons. The public sector and the government play multiple roles in society and the economy. They act as buyers of goods and services, suppliers of services, and regulators [2]. Public agencies provide the services and infrastructure necessary to sustain the social and economic structures in society.

Public procurement is typically divided into three phases-pre-tender, tender, and post-tender actions [8]. While a more detailed analysis recognizes six phases: (i) specification of needs, (ii) vendor selection, (iii) conclusion of contracts, (iv) ordering, (v) expediting, and (vi) evaluation and follow-up [20], we find the coarser-grained approach better suited for studying the state of practice in Finland. This preference is because the finer-grained phases are often internal to purchasing organizations, whereas our focus is on studying the tensions arising from stakeholder interactions overall.

There are several ways in which public procurement can occur. Firstly, before actual procurement, public agencies can collaborate with consulting vendors to prepare the tender, sometimes requiring a separate tendering process for this phase. The cooperation aims to establish a coherent view of the market, inform the market about the upcoming procurement, and communicate the requirements to the vendors participating in the tender. This collaboration is essential to plan and execute the process in a way that upholds the principles of non-discrimination and transparency [1].

Secondly, a supplier relationship is established through public procurement, mandated by legislation such as the Act on Public Procurement and Concession Contracts [13]. This relationship includes all vendors participating in public procurement, often forming a comprehensive ecosystem of companies. Finally, public agencies can purchase from in-house organizations, which the Public Procurement Act does not mandate. In-house procurement has unique characteristics because the procurement unit is not required to follow public procurement procedures, a significant deviation from the Public Procurement Act. However, in-house companies typically rely on public procurement when acquiring ICT services. Therefore, in-house companies have two roles as both a procurement unit and a service provider for public organizations.

Table 1. Interview Data

Organization	Id	Position	Field	Interview Duration
Procurement Unit 1	PU1	Chief position	ICT	47 min
Procurement Unit 2	PU2a	Manager position	ICT	48 min
Procurement Unit 2	PU2b	Senior Specialist	ICT	62 min
Procurement Unit 3	PU3a	Head of procurement	Procurement	63 min
Procurement Unit 3	PU3b	Manager position	ICT	49 min
Procurement Unit 4	PU4	Chief position	ICT	58 min
Procurement Unit 5	PU5	Manager position	ICT	56 min
Vendor 1	V1	Senior Principal	ICT	50 min
Vendor 2	V2a	Head of Department	ICT	49 min
Vendor 2	V2b	Specialist	ICT Procurement	49 min
Vendor 3	V3	Chief Position	ICT	45 min
Vendor 4	V4	Vice President	ICT Sales	56 min

3 Research Approach

Overall, ICT procurement as a human activity has received relatively little attention from researchers. Hence, the research questions we seek to answer are:

How do different stakeholder interpretations of public procurement regulation affect the ICT procurement?

We seek an answer via semi-structured interviews targeted at public organizations and vendors participating in public tendering. The semi-structured interview was the data collection method because it gives the best parts of structured and non-structured interviews [14]. The predefined structure guides the interviews with pre-formulated questions or themes, and all the interviews start with the same set of questions while allowing improvisation when needed.

Interviews were carried out and recorded between November 2021 and May 2022, and the details of individual interviews are listed in Table 1. The interview duration varied from 45 min to 63 min. The average duration was 51 min. Thematic analysis of the interviews revealed four themes related to public procurement norms, information systems, competence, and communication.

Procurement Unit 1 (PU1) is a government-owned enterprise (GOA), and its turnover is approximately EUR 140 million. Procurement Unit 2 (PU2) is a public administration with a budget of EUR 110 million. In the PU2, two interviews took place. In quotations, the separation between the two is marked with code PU2a and PU2b if necessary. Procurement Unit 3 (PU3) is a municipality with a yearly budget of EUR 740 million. PU3 had two interviewees, separated with abbreviations PU3a and PU3b if necessary. Procurement Unit 4 (PU4) is a city with a yearly budget of EUR 140 million. Procurement Unit 5 (PU5) has a

Table 2. Tensions in different ICT procurement phases summarized.

Pre-tender findings
Tension 1: Communication Between the Stakeholders.
Tension 2: Issues in Consulting the Vendors During the Process.
Tension 3: The Choice of ICT Procurement Opportunities and Resources
Tension 4: Invitation to Tender Has a High Impact on ICT Procurement.
Tension 5: Different Views on the Public Procurement Act.
Tension 6: Different Perceptions of the Suitable Solutions.
Tension 7: Attitudes Towards the Change.
Tension 8: Differences in Management Practices.
Tender findings
Tension 9: Most Advantageous Offer.
Tension 10: Purchasing Vast Systems Versus Purchasing Small Entities.
Tension 11: EA Management via Public Procurement.
Post-tender findings
Tension 12: Legislation Interfering with Stakeholder Relationships and Joint Roadmap.
Tension 13: Varying Methods to Manage Stakeholder Relationships.

yearly operating budget of EUR 375 million. PU5 is a joint municipal authority in the healthcare field.

Vendor 1 (V1) is an international ICT company. V1's turnover is approximately EUR 300 million, and V1 has 1100 employees in Finland. Vendor (V2) is an international ICT company with a turnover worth EUR 112 million and over 800 employees in Finland. Vendor 3 (V3) is a Finnish ICT company with a turnover worth EUR 42 million and approximately 500 employees. Vendor 4 (V4) is a Finnish ICT company. V4's turnover is EUR 2,7 million, and it has 23 employees.

4 Results

We have categorized the results to pre-tender, tender, and post-tender findings. These are summarized in Table 2.

4.1 Pre-tender Findings

Tension 1: Communication Between the Stakeholders. All the procurement units in this study employ preliminary market consultations with vendors and communicate with them during the pre-tender phase. These preliminary market consultations can take various forms. PU1, PU2, PU3, and PU4 consistently explore the market possibilities. Communication goes beyond formal connections with vendors via RFIs, although sometimes RFIs can be an excellent way to initiate a market dialogue with vendors. An RFI provides vendors

with an opportunity to inform the procurement unit about building new systems with modern technologies. For instance, as shown by V1, if the procurement unit is open to change and not overly tied to how the previous system functioned, an RFI can be a valuable tool for generating new ideas.

Another avenue for procurement units to familiarize themselves with market options is through everyday conversations and networking events with vendors. Vendors appreciate informal discussions because a better understanding of the procurement unit's needs often emerges from these interactions. PU1 and V2 highlight that when the procurement unit and vendor communicate openly, ICT procurement tends to be more successful. Similarly, V1 and PU1 emphasize that one of the least effective methods for acquiring an ICT system is to skip preliminary discussions with vendors and simply issue an RFQ. However, due to constraints like limited resources, time, and personnel, there are instances where ICT procurement may begin without prior communication.

Tension 2: Issues in Consulting the Vendors. Vendors believe that the procurement unit benefits the most from consultants' help if it can effectively communicate how it operates and what it aims to achieve. This allows the consulting vendor to understand the requirements for the new system better. V1 illustrates that some procurement units actively discuss options with other procurement units. For example, PU4 benchmarks and shares information with other municipalities about problems and solutions to find the most suitable option.

PU2 has had discussions within the organization about whether seeking consultation to prepare the RFP or RFQ is a part of the procurement process. Indeed, the Procurement Act [13] mandates preliminary market consultation, which is interpreted as a regulation for the pre-tender phase [8]. The Finnish Procurement Act states that preliminary market consultation with the vendor participating in the tender should not compromise the fairness of competition [13]. In the interviews, PU2a describes the approach as follows:

“Always before the tender phase, we review the familiar vendors, and, at the latest in the tender phase, we provide the opportunity for other vendors.”

Procurement units PU1, PU2a, PU2b, and PU3 acknowledge that in tendering, they need a clear understanding of procurement practices, and, as PU2a phrases it, “the game the vendors play.” It seems that this setting creates tensions regarding whether to trust that vendors prioritize the interests of the procurement unit or whether their incentives are misaligned.

Tension 3: The Choice of ICT Procurement Opportunities and Resources. In the pre-tender phase, public agencies also decide which opportunity to use for tendering. During the interviews, procurement units mentioned open, restricted, and competitive negotiated procedure opportunities for ICT procurement. When purchasing complex systems or something entirely new,

the competitive negotiated procedure often leads to the best outcomes. This procedure allows procurement units and vendors to communicate openly and comprehensively map out the system's long-term needs. For example, PU1 and PU2 use competitive negotiated procedures, typically resulting in favorable outcomes. However, PU2a believes that the competitive negotiated procedure can be demanding for the procurement unit, requiring resources such as expertise, time, and funds.

All procurement units agreed that direct awards are emergency solutions, often used in tandem with in-house purchases. PU3a and PU4 indicate that direct awards usually occur in vendor lock-in situations or when time is limited.

PU1 and PU3a emphasize that sometimes the legacy system must be replaced and included in the public procurement process, regardless of the high migration costs. PU2 believes that, in addition, the purpose is to respond to change proactively; sometimes, vendor lock-in can be calculated to be more beneficial for the procurement unit.

Tension 4: Invitation to Tender has a High Impact on ICT Procurement. Procurement units agree that the tender must be well-defined before publication and that errors are difficult to fix after the tender is public. PU3b says:

“Legal practice has proven that modifications are not allowed (in the tender), even if they are allowed in the law.”

Therefore, PU3b believes that the procurement practice needs revision. Before publishing the tender, the procurement unit should have a precise understanding of the expected outcome, even if it doesn't yet exist. The preliminary requirements must be adequate and precise because when the procurement unit receives bids from vendors, it needs to select the most suitable vendor based on the published criteria. In this phase, it doesn't matter if the procurement unit discovers flaws in the originally published tender because it cannot be modified. PU1 shares a similar perspective. PU1 criticizes the regulations for encouraging public organizations to rigidly follow procurement processes in environments that should be more adaptable towards agile methods.

Tension 5: Different views on Public Procurement Act. PU2b believes the procurement act enables free communication and agile development when used correctly. However, in Finland, the Procurement Act can be cumbersome for those who need to learn how to use it. On the other hand, PU1 suggests that the Procurement Act [13] encourages procurement units and vendors to engage in *“procurement theater”* where the procurement unit publicly carries out its legislative tasks, publishes RFP and RFQ, and receives bids from vendors. However, before this, the procurement unit has already selected the solution and the vendor. All the procurement units in this research acknowledge that there are occasions when they specifically require a certain product from the market. In practice, procurement units then define the requirements to align with only one vendor's solution or opt for in-house procurement.

Tension 6: Different Perceptions of the Suitable Solutions. The interviews revealed that vendors and procurement units want different things to some extent. As an example, procurement units in this research need ready-made systems. Purchasing SaaS solutions would be ideal. In addition, the Finnish government has given public organizations recommendations for cloud-computing systems.

In PU2, the organization's strategic objectives guide the planning of the software requirements in the tender phase. The top management has set the objective to refrain from purchasing customized products. In PU2, the minimum criteria for the software is that it has ready-made components and the user interface is modifiable. PU2a recons that the organization's IT landscape is complex and demands skillful personnel to manage it, and many times, the strategic skills to manage ICT procurement are missing.

In contrast, vendors' incentive is to offer tailored solutions for the procurement units, even if they can technically produce and deliver whatever is needed. V2, V3, and V4 all have similar messages on tailored systems, even though V4 plans to answer the market call in the future with a ready-made solution for case management.

PU5 recons that it is understandable if the procurement unit sometimes wants to acquire a tailored solution because the initial price is often tempting. However, tailored solutions carry great maintenance risks and may lead to vendor lock-in. In this research, procurement units and V1 depict that purchasing ready-made solutions is faster, easier, and more affordable than tailoring to procurement units' needs.

Tension 7: Attitudes Towards the Change. V1 and V4 point out that shifting the mindset in procurement units to adopt new systems and processes can be challenging. Many of these units have tailored their procedures to match the old system's performance, making it difficult to embrace change. For example, PU5 reveals that some Request for Proposals (RFPs) describe only the existing system's functions, limiting innovation. This rigidity in public organizations, as discussed by PU1, is often attributed to a lack of ambition to explore alternative work methods. V1 also suggests that public sector employees should take a more proactive role in implementing minor changes that can lead to significant improvements.

V1 and V2 highlight the presence of competent and innovative personnel in Finnish public organizations. However, their expertise remains underutilized due to daily job demands, leading to missed opportunities for enhancing processes and systems.

V4 emphasizes the success of small ICT entities, crediting innovative public sector leaders who have taken risks and embraced highly automated systems. The recurring message is that public organizations possess internal competence, which is not always harnessed optimally. The challenges of changing attitudes toward new systems and processes stress the importance of mindset shifts, leveraging existing competence, and fostering innovation.

Tension 8: Centralized Management Versus Decentralized Management. All public agencies in this study have multi-professional personnel responsible for publishing the RFIs, RFQs, and RFPs. The practicalities to take care of the procurement processes are centralized.

The procurement unit draws the initial requirements for the information system. Some procurement units, PU1, PU2, and PU3, have a project management office (PMO). In PMO, procurement units map out whether separate units in the organization have similar projects, if combining the resources is possible, and whether they have the resources to initiate the project. PU2 and PU3a depict that, at best, PMO processes enhance efficiency. PU1 has reduced all the duplicate ICT systems and vendors due to PMO functions.

PU1, PU2, and PU3 depict specialists from different units (business, ea, IT, procurement) evaluating their territory in PMO. Initially, the PMO scans the resources and determines whether the business case exists or initiates the project because the law mandates it. Naturally, the emphasis is on well-prepared projects and literature findings reveal that the RFQ requirements need to be carefully prepared because otherwise, the project may be prolonged, the budget may be exceeded, or the system may fail before production [4, 7–10]. Alarming, half of the procurement units in this research do not engage PMO practices and suffer from overlapping projects and systems.

4.2 Tender-Time Findings

Tension 9: The Most Advantageous Offer. The public procurement act in Finland [13] guides choosing the most advantageous offer, which often means the price has a heavy emphasis. PU1 says that the principle of enhancing the quality and lowering the price is flawed and unrealistic. PU1 has a strategy to set high basic requirements, ensuring that the participants' quality is good throughout the tender phase, and V2 has a similar idea. PU2a recons that the price is relatively demanding to erase from the selection criteria even if they have tried. Many vendors can meet the initial criteria; only price matters after that.

PU2a depicts that for some, it is demanding to calculate the most advantageous offer. PU2 has learned from experience how to calculate and estimate lifespan costs. PU4 depicts similarly; experience helps to scan the apparent pitfalls in planning the system, procurement, criteria, and vendor selection. PU1, PU2, and PU4 are wise to interview the vendor's team and set soft criteria such as the team's vision, competence, and ambition to make the best vendor decision. Thus, more than merely defining software requirements and the price is required in ICT procurement. However, procurement units need help to implement soft criteria in the selection criteria because the overall price for a good team is demanding to evaluate.

Tension 10: Purchasing Vast Systems Versus Purchasing Small Entities. PU3a recognizes two main methods to build the tender. Sometimes, PU3 purchases the platform and the development in one RFQ, and sometimes, everything is purchased separately: platform, development, and maintenance. PU1

and PU2 emphasize that the entities they wish to purchase need to be appropriately sized - the too vast a system is demanding to manage and causes vendor lock-in. However, all the procurement units recognize that stakeholder management becomes complex if the number of vendors rises, and procurement units hope for top-down support. V4 depicts that the requirements are the same for small and large public organizations because they are under the same legislation. For example, small and larger municipalities need similar governance and case management accuracy. V1 thinks similarly that public organizations waste resources to define requirements for the new ICT system because other public organizations have usually tackled the same issue.

Tension 11: EA Management via Public Procurement. Enterprise architecture (EA) management via public procurement is challenging. PU1 and PU2 reckon that vendors may not be interested in planning the solutions to fit the existing EA. PU1 hopes the vendors will adopt a holistic view of the buyer's EA when the same vendor provides different solutions to different procurement units in the same public organization.

Currently, procurement units depict that the vendors are only sometimes invested in taking the time to familiarize themselves with public organizations' existing operations and systems. PU2 reckons that smaller vendors are more interested in delivering easily deployable and manageable solutions and are more flexible than the larger vendors. Migration costs can increase if the existing EA is outside the selection criteria. PU2a thinks that PU2 is a more significant customer to the small vendors than to the large vendors. As a small business, V4 agrees with the view.

The procurement unit's EA has varying ways to emerge in the tender phase. PU1 field of business is mission-critical; software-wise, everything they purchase must go through many official checks. PU1 manages the tendering practices top-down; procurement units cannot solely purchase something that fits their purposes. The purchasing practices support standardized technology solutions and sustainable software lifespan management.

PU2 uses the JHS-179 standard to define the target architecture to avoid surprises in the implementation [18]. Furthermore, in PU2, IT governance sets objectives for the tender. In the tender phase, PU4 describes the current state of EA. In addition, PA4 describes the target stage EA in advantaged ICT procurement. Like PU4, PU3 uses the current state EA descriptions in the tender phase. PU5 depicts that the organization's EA does not show in the tender. Usually, EA is examined after the vendor selection in the post-tender phase, which is costly, complex, and prolongs the project. PU5 describes that the current EA initiatives exist but do not show in practice.

4.3 Post-tender Findings

Tension 12: Legislation May Interfere with Stakeholder Relationships. Public procurement legislation may interfere with prosperous stakeholder relationships, so procurement units reckon it would be convenient to predict future needs in the tender phase. Essential changes are impossible during the contract

and may lead to vendor change. PU1 depicts that sometimes they have flourishing cooperation with the vendor, but the law causes unnecessary vendor changes. For example, the original software works well, but a new need emerges near the original solution. It could be effortlessly developed with the existing vendor, but the public procurement act in Finland does not allow essential changes in the contract period [13]. As a consequence, new procurement needs to be initiated.

PU2a says that sometimes they try to include consulting services in the RFQ and demand that the solution be used in all procurement units to avoid the abovementioned issue. However, PU3b sees pitfalls in this approach. Even if the solution could be used in the other procurement units, the price is considered an essential change, which usually demands the beginning of the new procurement process. In addition, PU2a realizes that the tactic is only sometimes successful because future needs are almost impossible to predict.

PU1 and PU2b emphasize that the more important thing is to keep an excellent record of stages, development, and tasks if the vendor changes due to legislative or other reasons. When the existing system works well and the stakeholder relationship is good, changing the vendor and system wastes resources for the procurement units. In this research, V4 depicts that they wish to produce their services so that the procurement unit never suffers vendor lock-in with them. Instead, they wish to continue cooperation because it has been successful.

Tension 13: Methods to Manage Stakeholder Relationships Vary. Traditionally, public agencies have paid the vendors in installments, and if they disagree with the performance, they may refuse to pay the installment. Another way to manage the contract period is to set vendor fines. Furthermore, some agencies use the option to continue the vendor contract for the next period as a carrot. PU1, PU2, and PU3 reckon these methods encourage rigid and waterfall-like software development. Furthermore, PU1, PU2, and PU3 reckon that the vendor should be ambitious to produce its services with quality rather than be pressured with installments and fines to produce barely acceptable services and products. PU2b thinks it is within the procurement unit's management culture whether they can motivate vendors without using ramifications. V1 and V4 depict similarly but from different points of view: attitude and ambition need to be towards solving problems together and offering the best possible solutions for the procurement units.

In Finland, in-house procurement is a rather significant phenomenon. In in-house purchases, PU4 thinks the installment with-holding is the only option to receive acceptable solutions. In-house procurement is considered a part of the procurement unit's internal production even if the decision-making and governance are separate, which may cause an issue in quality control. PU3b thinks the permanent contract motivates the vendors compared to the temporary contract with the option for the second contract period. The assumption is that the vendor appreciates continuous and good business relationships as much as the procurement unit. PU1 depicts that they use service level agreements (SLAs) in the contract period, which could be better. All-in-all, procurement units in this study agree that the public sector uses far more sticks than carrots in vendor relationships, which does not work.

	Public Sector	Vendors
Public Sector	<ul style="list-style-type: none"> Procurement units have varying perceptions of norms (Tension 5): <ol style="list-style-type: none"> Norms do not obstruct the tender phase, but the ability to leverage legislation differs. (Tension 7). Norms steer towards a waterfall system development and impede innovation. (Tension 3 and 4) The principle of reducing costs while enhancing quality is flawed (Tension 9). Some procurement units utilize software reusability contracts and decentralized PMO. (Tension 8) Purchasing practices vary between large systems and small entities (Tension 10). Drafting tenders carefully is crucial because they have a significant impact on ICT procurement outcomes (Tension 4). 	<ul style="list-style-type: none"> Norms result in unnecessary vendor changes during the post-tender phase (Tension 12). Direct awards and in-house procurement lead to extended projects and high migration costs (Tension 5, Tension 9). There is an incentive to acquire ready-made solutions (Tension 6, Tension 11). Vendors often show little interest in the existing enterprise architecture and system interoperability of public organizations during the tender phase (Tension 11). Including soft criteria in the selection process can be challenging (Tension 4, Tension 5). Unofficial communication sometimes results in vendor selection before the tender phase, which may lead to legal consequences (Tension 1 and 5).
Vendors	<ul style="list-style-type: none"> Norms do not impede effective practices (Tension 7). The choice of ICT opportunities impacts efficiency, vendor selection, and the success of ICT procurement (Tension 2, Tension 4). The public sector lacks ambition in developing practices (Tension 7, Tension 8, Tension 13). Procurement units are slow to change ideas, processes, and practices (Tension 7, Tension 8). 	<ul style="list-style-type: none"> Vendors have different approaches to what they offer (Tension 1 and 6): <ul style="list-style-type: none"> Tailored systems to secure an irreplaceable position. Easily deployable and replaceable entities.

Fig. 1. Public Sector and Vendor Relationships

5 Discussion

5.1 Research Questions Revisited

This research studies how different interpretations of regulatory aspects affect public ICT procurement. We identified 13 tensions in ICT procurement, which fit into four categories. Figure 1 summarizes public sector and vendor relationships in detail and describes where the tensions arise. Below, we list some differences in interpretations that contribute to the tensions.

- *Public Procurement Norms*: Public ICT procurement is highly regulated and normative. However, while the norms set the field for the processes and stakeholders, the human aspect is vital. The perception of public procurement norms raises tensions.
- *System*: In ICT procurement, the system or service is at the center of the acquisition. However, interestingly, the procurement units and vendors depict that the technology does not hinder finding efficient and well-functioning solutions. Different intentions, ideas, and ambitions are the most significant obstacles. Vendors and procurement units want to acquire and provide different solutions, and procurement units prefer ready-made solutions, whereas vendors are incentivized to offer tailored systems.
- *Competence*: In some procurement units, the quality aspect is strong, and these public agencies aim to reduce the price’s effect on the selection criteria. Pre-tender phase and preliminary market consultation are critical in such evaluation. Furthermore, certain aspects, such as the vendor’s ambition, the team’s competence, and vision, are challenging to put in the selection criteria. Therefore, suppose the preliminary market consultation reveals the most suitable option, which is not the cheapest.

- *Communication*: Communication with vendors – unofficial conversations, preliminary market consultation, and bench-marking – is vital for the procurement units before publishing the tender. The tender has a high impact on ICT procurement because it affects vendor selection, system requirements, and interoperability, duration of the project, and efficient use of resources. In addition, carelessly drafted RFP or RFQ may lead to legal ramifications. Drafting the tender is particularly demanding for the procurement units because errors are almost impossible to correct after publication. System requirements and interoperability must be included in the tender because the vendor is selected against these criteria.

However, in practice, all procurement units recognize that sometimes vendor selection happens before the tender phase, even if the incentive in law is to ensure fair and equal competition. The communication between the procurement unit and vendor is regulated, especially in the tender phase and preliminary market consultation [13]. Both parties, procurement units, and vendors realize that the Public Procurement Act guides communication for a reason. However, the balance between open communication and favoring should be found simultaneously.

5.2 Threats to Validity

In this paper, five procurement units and four vendors participated, and twelve interviews were done. The research method, semi-structured interviews, allowed the interviewees to depict what was significant to them. However, this might be a weakness as well. Semi-structured interviews combine parts from structured and non-structured interviews [14], and eventual consistency comes from the pre-selected themes and the freedom to specify and elaborate on subjects that emerge during the interviews. Hence, the research method fits the study, contributing to the research approach's validity.

Data is collected and analyzed systematically, in an iterative way, and rigorously, which increases reliability. However, the sample size introduces some issues of generalisability [15]. Another issue related to the sample is that they all are from Finland, so results cannot be generalized to other countries due to differences in national legislation. However, the procurement units and vendors in this research cooperate and, in some cases, depict their relationship. Therefore, the consistency in results and similar findings in the literature reveal that the study has validity even if the sample size is small [15]. Hence, even if the sample size prevents the final conclusion on the subject, the results are significant to share with the research community.

Finally, inner validity could be improved with triangulation or multiple researcher evaluation [14]. Here, the authors directly make deductions that may infer the inner validity. However, the results and the deductions have been reviewed and accepted by independent inspectors in the thesis process; this work is based on [omitted-for-blind-review].

5.3 Future Work

Public procurement issues are recognized in literature and practice. However, public procurement is a separate regulated process in literature rather than a part of the communication and cooperation of humans, which will be fundamentally required to complete a procurement. Closing this research gap is a part of our future work, even if this research is a significant initial step. Hence, holistic exploration of ICT procurement is a vital topic to cover.

Procurement units in this study recon that it is almost impossible to predict all future needs, and they prefer exit points if the vendor relationship becomes challenging. Hence, the post-tender phase concerning the agility to change vendors would be interesting to cover. In some interviews carried out in this research, the in-house purchases caused issues. In-house procurement is not within the procurement regulation, which for the cooperation does not follow the standard practices that apply to vendors. The regulatory aim is to enhance efficiency in public procurement. These two aspects hinder effective practices in this study.

6 Conclusions

In this paper, we studied how procurement units acquire software. Based on semi-structured interviews, it was found that the agencies have different interpretations of the Public Procurement Act [13]. In light of the Public Procurement Act, a durable vendor relationship is challenging to establish. Hence, careful project preparation is vital in public procurement; considering the entire software lifespan needs in one tender could be helpful in practice. Moreover, decisions on how and what entities to purchase must be well thought through. Procurement units and vendors recommend tracking the ICT procurement process and system development to facilitate vendor change if it is needed when something essential changes.

References

1. Directive 2014/24/EU of the European Parliament and of the council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02014L0024-20180101>
2. Astbrink, G., Tibben, W.: ICT accessibility criteria in public procurement in oecd countries – the current situation. In: Stephanidis, C., Antona, M. (eds.) *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion: 7th International Conference, UAHCI 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I*, pp. 155–164. Springer Berlin Heidelberg, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39188-0_17
3. Baily, P., Farmer, D., Crocker, B., Jessop, D., Jones, D.: *Procurement principles and management*. Pearson Education (2008)
4. Baily, P., Farmer, D., Crocker, B., Jessop, D., Jones, D.: *Procurement Principles and Management*. Pearson Education Limited, Edinburgh Gate. Harlow. Essex CM20 2 JE. England. (2008)

5. Ghezzi, R.K., Korhonen, M., Vilpponen, H., Mikkonen, T.: The role of in-house procurement according to Finnish municipalities' purchase invoice data. arXiv preprint [arXiv:2211.14570](https://arxiv.org/abs/2211.14570) (2022)
6. Guzmán, F., Sierra, V.: Public-private collaborations: branded public services? *European Journal of Marketing* (2012)
7. Holma, A.M., Sammalmaa, J.: Julkisen hankintaprosessin alkuvaiheet hankintatoimen ja toimittajien välinen vuorovaikutus (in Finnish). University of Vaasa (2018). <https://urn.fi/URN:ISBN:978-952-476-843-6>
8. Holma, A.M., Vesalainen, J., Söderman, A., Sammalmaa, J.: Service specification in pre-tender phase of public procurement - a triadic model of meaningful involvement. *J. Purchas. Supply Manage.* **26**(1) (2020). <https://doi.org/10.1016/j.pursup.2019.100580>
9. Iloranta, K., Pajunen-Muhonen, H.: Overview of Enterprise Architecture work in 15 Countries. Tietosanoma, Helsinki (2012)
10. Keränen, O.: Roles for developing public-private partnerships in centralized public procurement. *Ind. Mark. Manage.* **62**, 199–210 (2017)
11. Kopponen, A., et al.: Empowering citizens with digital twins: a blueprint. *IEEE Internet Computing* (2022)
12. Koski, A.: On the Provisioning of Mission Critical Information Systems based on Public Tenders. Doctoral Dissertation, University of Helsinki, Finland (2019)
13. Ministry of Employment and the Economy: Act on Public Procurement and Concession Contracts. Ministry of Economic Affairs and Employment, Finland (2016). <https://www.finlex.fi/fi/laki/kaannokset/2016/en20161397>
14. Myers, M.: *Qualitative Research in Business and Management*. SAGE Publications (2020)
15. Noble, H., Smith, J.: Issues of validity and reliability in qualitative research. *Evid. Based Nurs.* **18**(2), 34–35 (2015). <https://doi.org/10.1136/eb-2015-102054>
16. Nuottila, J., Aaltonen, K., Kujala, J.: Challenges of adopting agile methods in a public organization. *Int. J. Inf. Syst. Proj. Manag.* **4**(3), 65–85 (2016)
17. Stilger, P.S., Siderius, J., Van Raaij, E.M.: A comparative study of formulas for choosing the economically most advantageous tender. *J. Public Procurement* **17**(1), 89–125 (2017)
18. Suomidigi: Design and development of the enterprise architecture for public sector. <https://www.suomidigi.fi/ohjeet-ja-tuki/jhs-suositukset/jhs-179-kokonaisarkkitehtuurin-suunnittelu-ja-kehittaminen> (In Finnish) (2017)
19. Tahir, M., Vaheed, M., Burhanuddin, M.: Theoretical foundation in analyzing gaps in information communication and technology (ict) tender process in public sector. *J. Eng. Appl. Sci.* **13**, 1407–1413 (01 2018). <https://doi.org/10.3923/jeasci.2018.1407.1413>
20. Van Weele, A.: *Purchasing and Supply Chain Management*. UK, Cengage Learning EMEA (2018)



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Improving Communication and Collaboration in Enterprise Architecture Projects: Three Propositions from Three Public Sector EA Projects

Ari Rouvari^(✉)  and Samuli Pekkola 

Faculty of IT, University of Jyväskylä, Jyväskylä, Finland
ari.rouvari@tuni.fi

Abstract. Enterprise architecture (EA) is infamous for implementation problems and unredeemed promises. Imprecise and unstandardized EA work practices and various definitions make it difficult to comprehend what should be done and how, and to advance digital transformation. Earlier studies have identified communication and collaboration challenges as one of the most common and fatal sources of problems. In this paper, we study how different actions help avoiding and addressing communication and collaboration problems in EA projects. We conduct a qualitative and comparative case study of three public sector EA projects in Finland. Our data is based on ethnographic observations, which were later inductively analyzed. As an outcome, we present a theoretical explanation of the phenomenon and make three propositions to manage and possibly overcome the problem.

Keywords: Enterprise architecture work · public sector · communication and collaboration · problem · qualitative case study · ethnographic approach

1 Introduction

Organizations are investing in digital transformation and creating accessible digital services [14, 15, 38]. In this context, enterprise architecture (EA), an information management tool that helps them visualize and execute their strategies, describes the strategy, business, data, apps, and technology architectures and connections between them. EA is an appropriate method and has an important strategic and operative role in the digital transformation of organizations and ecosystems [23, 28, 35]. As a tool for managing their digital transformation processes, EA helps to create new digital capabilities and service ecosystem culture.

EA implementation and utilization projects are infamous for their problems [13, 39]. The most common issue is collaboration and communication among different partners and stakeholders [7, 37]. Earlier recommendations to solve the problems are impractical since the suggestions are rather generic [13, 20, 39], while EA problems are highly contextual [37]. There is thus a knowledge gap on how to cope with the communication

and collaboration problems in the EA projects. This motivates our research. We seek answers to: How can communication and collaboration problems in EA projects be addressed? What consequences are expected from these activities?

We conducted a qualitative and comparative case study on three large-scale digital transformation projects utilizing the EA approach in the Finnish public sector. We wanted to understand how the EA project owners and team members address emerging communication and collaboration problems through different actions. We also studied the impacts of those actions. We constructed a simple model and used it to analyze the data from ethnographic observations. We argue that the communication and collaboration problems can be mitigated even during the projects by increasing and reallocating resources or changing the working practices. It requires sensitivity and distance to identify them and authority to change the situation.

The paper is organized as follows. First related research is summarized. That is followed by the research settings and methods section and our findings. The paper ends with a discussion and conclusion sections.

2 Related Literature

Digital transformation is about digitalizing the organization's services, functions, processes, and transactions. EA is a holistic approach to helping digital transformation by illustrating various details and their relationships, handling communication issues, understanding business needs, and addressing complexity and integration issues [10, 16, 30]. Social and organizational challenges and unexpected incidents impact intense digital transformation [1, 15, 42]. EA is an information management tool, and it can be used for organizations' management for different purposes [24].

EA aims to provide a holistic view of the organization and its business, data management, applications and technologies, their current and future states, and how to reach the goals [22, 41]. It will benefit organizations if they achieve various dynamic EA capabilities [2, 45]. High-quality EA is defined through seven quality attributes: alignment and integrity, the quality of EA products and services, maintainability and portability, scalability, security, reliability, and reusability [32].

EA projects tend to be large and complex. They bridge multiple departments and levels and have myriad stakeholders and several viewpoints, which make them failure-prone [13]. These failures have been studied, for example, in the public sector in general [13, 29], in government agencies, municipalities, and higher education institutions [39], and in many other settings e.g. [3, 31]. The challenges are usually not technical but relate to leadership, governance, management, staff commitment, and governmental politics [5, 21, 22]. Kaisler et al. [22], for example, recognized communication challenges between middle management, managers, and other EA stakeholders, especially on methodology and modeling issues. The problems correlate and are interwoven in convoluted causal chains [18], which makes the situation even more complex.

EA management challenges are related to EA documentation, EA planning, and EA communication and support [11]. EA project challenges are associated with the EA definition and documentation, flexibility, time pressure, and complexity [33]. The biggest challenge of the EA practices is communication between decision-makers and

stakeholders [25]. As EA development is mainly about communicating and collaborating with different stakeholders, the problems there escalate quickly and cause severe issues in EA projects. Communication and collaboration problems have been identified as being common in EA projects, which also explains other EA obstacles [7]. As communication and collaboration are influenced by twenty factors [6], ranging from technical to organizational and personal issues, solving them is not easy. However, it is vital for the EA projects as they are a means for engaging the stakeholders [27], especially when they have varying backgrounds and experiences [12].

In these situations, EA artifacts, models, and descriptions are used as a communication tool [34, 44]. This, in theory, solves some of the communication problems as the models provide a common point of reference and a common language [26, 34]. Similarly, different statements have been made about paying attention to success factors and problematic issues [13, 20, 39]. Even the importance of communication skills has been acknowledged [46]. Yet, the communication problems and failing EA projects persist. One of the reasons is the context specificity of the EA and EA projects [17, 46]. Especially communication and collaboration are highly contextual and temporal [21, 37].

3 Research Methods and Settings

To understand how communication and collaboration problems are addressed in the EA projects, we conducted a qualitative and comparative case study on three public sector EA projects in Finland (c.f. [47]). We paid attention to communication problems and their root causes, to actions to solve them, and to those actions' possible implications.

We derived the data from the first author's retrospective analysis of his ongoing EA projects. He has been working for more than fifteen years as a chief enterprise architect or consultant in numerous EA projects, mainly in the public sector. For this study, we chose his three recent EA projects where communication and collaboration challenges have been identified as critical. As he has been actively involved in the projects, he had a unique chance to gain in-depth data and understanding about the projects, their challenges, and actions. In this paper, we rely on his ethnographic fieldwork e.g. [36], and project documentation, such as memos, project plans, and meeting minutes. With ethnographic observations, contrary to action research [8], where the researcher aims to change the situation, the researcher solely observes and reflects on different situations and actions. Although we were interested in corrective actions to solve the challenges, the first author was not in a position to actively pursue their solving – being an architect or consultant, one can merely inform the project owners about the challenges and hope for the best. There was very little he could do.

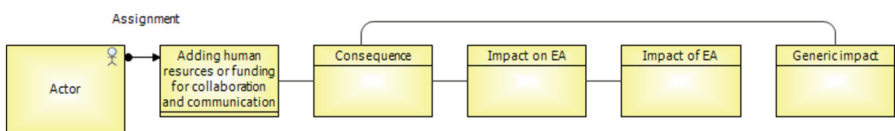


Fig. 1. The model for data analysis.

To structure our analysis, we used a simple model influenced by the activity theory [9] (Fig. 1). The actor, an individual or a community, does an action. An action has one or more consequences (outcomes) that affect the EA (impact on EA). The EA continues to impact, for example, the development of its domain (impact of EA). Generic impacts are the aftermaths of all these.

Our data analysis proceeds as follows. First, the first author identified and analysed the communication and collaboration problems on two different occasions: in winter 2021 and in summer 2022. Although he was aware of various classifications, the analysis was data-driven and inductive. He classified the problems as critical (the situation is chaotic, elevation is unlikely), challenging (the situation is challenging but solvable), or desirable by using his experience as an actor in these projects. He then wrote an anonymized storyline of each project and its activities. These storylines and the first author's experiences were used in the structured analysis of each project. Finally, the analyses were merged to create a more generic theoretical model. Although the first author analyzed the data, the findings were constantly discussed among the authors to reduce potential single-researcher bias.

Next, we will present each EA project, its storylines, and the impact chains.

4 The Cases and Observations

In this section, we present our analysis of three EA projects.

4.1 Project A

Project A is a national reference architecture by a Finnish government agency. The EA development started in Q3/2019. EA project described the baseline and target stage architectures, which include 78 strategy, business, data, and application architecture artifacts (65 diagrams and 13 tables). The architecture is already published. Initially, the project had four stakeholders and an architecture team of five members. By Q1/2022, the number of EA team members has more than doubled, and the number of stakeholders has increased by two new organizations.

In winter 2021, communication and collaboration challenges were severe as the EA team had only one EA consultant (the first author) and some representatives from Government Agency A. In summer 2022, the situation improved because the owner of the EA project increased the project's human resources and intensified communication with the domain agencies by surveying to check whether the architecture was understandable and correct.

In early 2021, a new enterprise architect and a domain expert from an agency joined the EA team. They aimed to improve the EA work and bring in necessary competencies. This had positive impacts on the EA: the EA method was used better, and the quality of the EA artifacts improved. They became more understandable and usable. The evaluation survey focused on the architecture documents. The six reviewers felt that various items were comprehensively described, but also made suggestions for improvements, many of which were noted, fostering the rigor and accuracy of the architecture. In addition, the project owner (Government Agency A) uniting two similar EA projects from neighboring

domains with many links and forms of collaboration. Figure 2 shows how the reallocation of resources, in this case merging two EA projects (action), improved understandability (consequence), harmonized the EA definitions (impact on the EA) and improved their interoperability (impact of the EA).

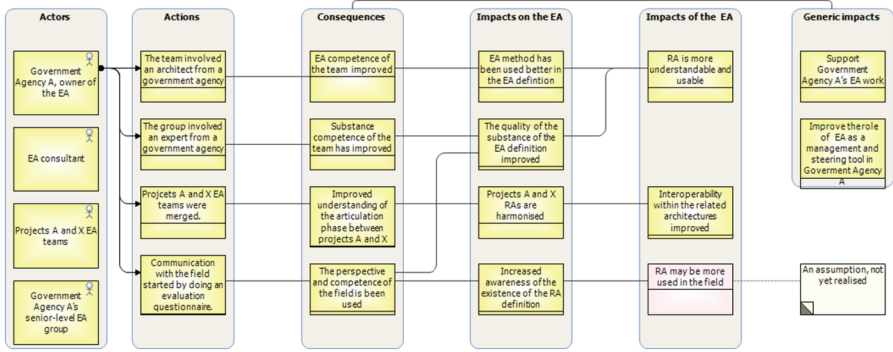


Fig. 2. Detailed actions and impacts in Project A.

Another means to improve communication and collaboration was the earlier-mentioned survey to assess the unambiguity and clearness of the EA definitions, identify shortcomings, and suggest improvements. It was conducted in parallel with the contiguous EA projects. It received a positive response and helped to improve the EA definitions. In other words, the survey increased general awareness of EA, domain knowledge, EA quality, and EA artifacts fit with the practice and practical needs.

There were two generic impacts: the actions and their consequences supported Government Agency A’s EA work and improved the role of EA as a management and steering tool.

These improvements can be explained by the increment of the EA team membership. In three years, the project more than doubled the number of architects and specialists, which provided adequate resources and skills to EA artifact development and cooperation and dialogue with the government agency and other stakeholders. They became aware of how critical communication and collaboration are in the EA projects. One of the project’s success factors was simply the increase of resources.

4.2 Project B

Project B is a national reference architecture owned by the same government agency as in Project A. Its descriptions focus solely on the target stage architecture and strategy, business, data, and application descriptions. The architecture consists of 82 artifacts (58 diagrams and 24 tables) In Q1/2021, the project had three stakeholders and an architecture team of seven members. In Q2/2022, the situation changed significantly when Government Agency A launched a new extension project involving 29 new organization members and more than 100 new strategy experts, architects, and other specialists. This extension project continued and replaced the first project. The main driver for launching

the extension project was to improve communication and collaboration within the field since this was found problematic in the first project.

In winter 2021, the lack of communication and collaboration had become critical because the EA team had only one EA consultant (the first author) and two representatives from the agency. By summer 2022, the situation had improved due to several actions taken within the year. First, another architect and an agency CIO were invited to join the EA team. Some domain experts and technical specialists were encouraged to attend the meetings, which increased EA and domain competencies and provided better awareness and understanding of the target area. It further influenced the EA artifacts and their quality and applicability in the domain and the use of the EA method in general. Second, Government Agency A aimed for better inter-organizational collaboration in the public sector. The Finnish public sector has traditionally been organized into sectors, each responsible for its area and tasks. The agency tried to break these siloes by encouraging, enforcing, and funding collaboration – and using EA to achieve this. This new EA project aimed to develop the reference architecture with a diverse group of representatives. Thus, a large number of organizations joined the project. It had three-fold implications: it increased the awareness of the current reference architecture descriptions, improved the quality of the EA artifacts, and made future reference architecture implementation much easier. As a result of the actions and their consequence and impacts on EA, we assume that stakeholders will have better opportunities to achieve the project objectives.

These actions, consequences, their impacts on EA, and impacts of the EA will improve the EA’s role as a management and steering tool for Government Agency B. Also, collaboration and EA work will be more effective as a good example is provided. Figure 3 shows this impact chain: how adding another architect to the EA team (action) improved the team’s competence (consequence), resulting in the EA method (impact on the EA) and the better usability of reference architecture (impact of the EA).

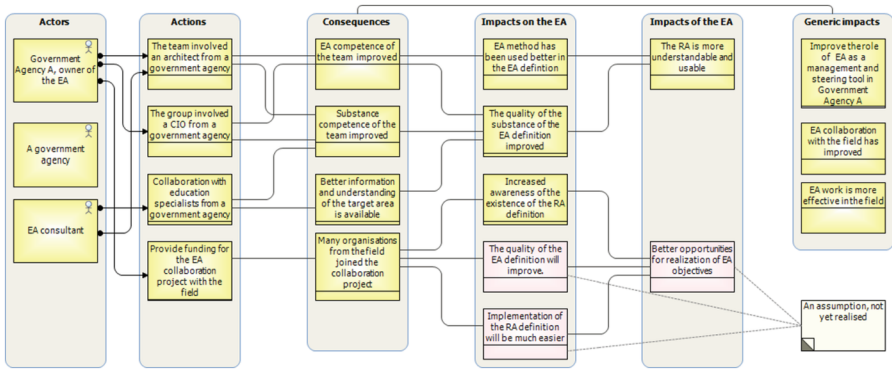


Fig. 3. Actions and impacts in Project B.

Project B illustrates the power of corrective activities during the project. Almost right after the start, the project faced several communication and collaboration challenges. These were solved immediately and significantly investing in human resources in the project. As the team was then able to provide benefits, some concrete, some potential,

Government Agency A decided to fund a new two-year follow-up collaboration project, replacing and continuing the first one. The new project involves 29 new organization members and more than 100 employees.

4.3 Project C

Project C is a national enterprise architecture owned by another Finnish government agency. It started Q1/2019 and closed Q2/2022. The project aimed to develop an EA architecture for a new government agency. The architecture focused on the target stage descriptions and included strategy, business, data, and application architectures. It had 105 artifacts (86 diagrams and 19 tables), all published. The project had four stakeholders and an architecture team of six members.

In the project, the EA team felt severe collaboration and communication problems with their stakeholders and owners. The EA team was thus active and pushed the agency to collaborate and arrange meetings to improve the EA and its interoperability with their other architectures. This push and these meetings improved semantic and technical interoperability between the architectures. Ultimately, in the future, this capability will hopefully deploy to different services between the agencies.

Government Agency B meetings increased confidence in the EA team: as a result, the agency representatives gave some extra tasks to the EA team. The team also marketed EA actively, further increasing the awareness of their work. These actions increased the EA team's motivation, influencing the quality of the EA artifacts.

However, the situation did not proceed smoothly. Due to the personnel changes in Government Agency B, one of the related architecture projects was halted and not published, which jeopardized the interoperability of the architectures because the relations and the responsibilities had to be reconsidered.

Another change took place when a lawyer from Government Agency B joined the EA team, which increased the team's motivation. They were able to create new EA artifacts where the forthcoming legislation was understood and incorporated. The relationship had mutual benefits as the lawyer better understood the boundaries set by the EA and was able to consider those when writing the legislation proposal.

The EA team also participated in the agency's strategy process. Constant criticism and debate whether the proposed new organizational structure was needed however, created frustration among the EA team members. Luckily, this did not affect the EA descriptions, only communication with other stakeholders.

The EA team hired some external help. They contracted an experienced external enterprise architect from the same domain to evaluate the artifacts and elaborate on some project details with the team. The team was thus keen to improve the EA and ensure that it is understandable and usable by all parties. As a result of this mini-evaluation, the business model view was added to the EA artifact. It will thus contribute better to the new agency and its future operations.

The estimated and already experienced success of the EA project motivated the EA team members and their work in their home organizations. The project will have far-reaching impacts beyond a single project. Figure 4 shows how the lawyer's joining the EA team (action) motivated the team (consequence). The legal capability impacted the

EA definition by improving its legal interoperability. On the other hand, the EA work supported the writing of the act (impact of the EA).

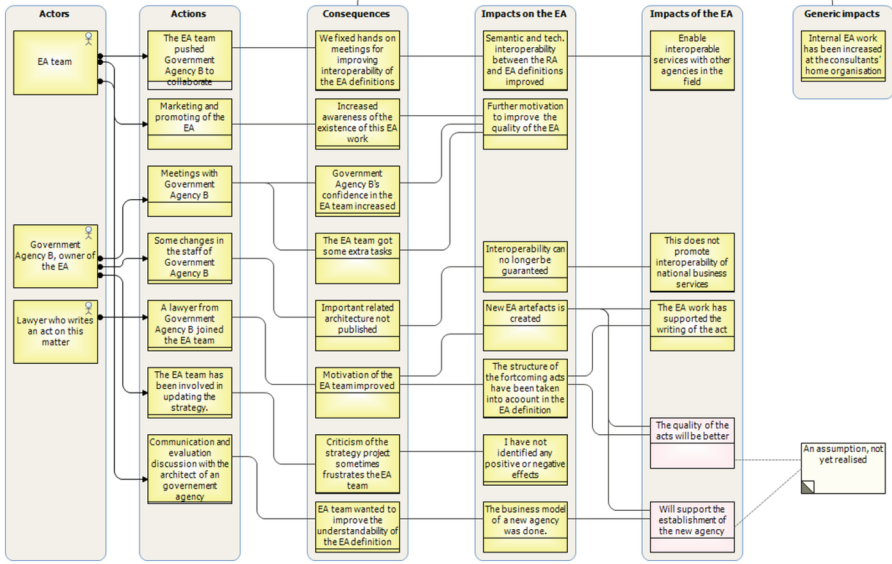


Fig. 4. Actions and impacts in Project C.

In Project C, the EA team was balanced and efficient in their actions. Each member had a specific role and responsibilities. They worked well, were motivated, and actively sought solutions. The activities were visible and appreciated. It is illustrated by a lawyer from Government Agency B who joined the group – she perceived the team supported her in writing a new law – and by participation in the agency’s strategy process.

5 Discussion

Our cases demonstrate that collaboration and communication can be improved by either reallocating the resources, changing the ways of working, or both. However, these activities usually require top management’s support or decision. It follows that it is essential to increase the awareness and knowledge of EA among senior management. In this endeavor, the enterprise architects’ communication and leadership skills are emphasized [18]. The owner of the EA project may, like in all our projects, add resources, such as people, money, or technologies, to the project to boost collaboration. On the other hand, as Project B illustrates, the EA team can improve collaboration by tuning the way they work and rearranging work processes – even during the project. Supplementary architecture descriptions and domain-related competencies from other government agencies improved cooperation between Government Agency B and the government agency, which, with enhanced working processes, fostered the EA team’s architecture capability maturity and efficiency. When these were further reflected in the project results, the

architecture definitions and EA artifacts quality improved, making them rigorous and accurate. The architecture descriptions and documents are consequently executable and, for example, more interoperable with related architectures.

However, the owner's actions may easily hinder or destroy such progress. In Project A, the project owner changed, and new priorities were introduced, which slowed the progress. In Project C, a related EA project was terminated, so Project C had to be re-scoped and replanned. Interoperability issues are thus compromised when related architectures are not published or the projects face challenges. Here, the role of the project owner is critical: if she is not satisfied with the actions and progress of EA work or the EA team members, the changes are evident. Due to the multiple connotations of EA work [34], such frustrations and displeasures emerge unchallengingly. They emphasize the collaboration and communication skills of EA teams [46].

Figure 5 summarizes all three cases and generalizes our observations. The main actors are the EA project owner and the EA team taking the actions, while external, reallocated resources (such as a lawyer in Project C) may also influence them. The main actions to be executed are reallocating resources or changing the working methods. They increase the EA team's competencies in actual EA work and communicate and collaborate with others. It, in turn, improves the quality of EA work and artifacts and furthers their usefulness.

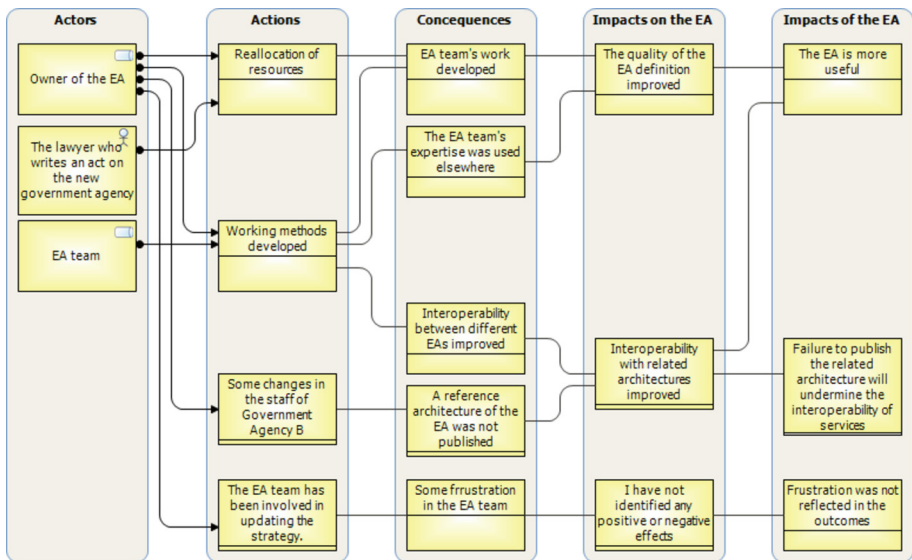


Fig. 5. Actions and impacts on the lack of communication and collaboration in Projects A, B and C.

Despite the conditions and contexts and their influence on EA management [4, 17], we abstracted the contextual-specific communication and collaboration problems from three public sector projects to general actions and impacts. From these, we derive three propositions for EA project practitioners to prevent the obstacles.

Proposition 1: In EA projects, management can improve communication and collaboration by reallocating resources in a controlled manner.

This proposition is in line with [2] that human EA management resources have a strong influence on the development of EA management. It is in line with the observations EA has problems with gaining the project management's commitment [5]. Even the architects need organizational and executive support and adequate resources [21].

In Project C, the EA team was invited to participate in the strategy work, but conflicting expectations emerged. All stakeholders were not committed to a common goal. One member of the strategy team even considered the whole strategy pointless. It demotivated the EA team and undermined their work. These conflicting priorities and the absence of the stakeholders' shared view are typical engagement problems in EA [27]. Under the circumstances, collaboration is challenging to improve by increasing communication or resources if there is no shared goal. Such a lack of stakeholder involvement causes several other problems [18].

In Project A, increasing the project's human resources and conducting a survey solved many collaboration and communication problems. However, resource reallocation also created new challenges when the team's way of working changed. Similarly, Project C faced new challenges. It means that collaboration and communication must be taken into account in the EA project plans as they likely influence how the resources can be used. During the project planning phase, the key stakeholders need to be identified, and the different forms of collaboration and communication need to be planned and documented. Corrective actions, like in Project B, may not always be identified or appropriately executed. The lack of collaboration and communication must thus be considered similarly to any potential risk and addressed in the risk assessment and mitigation plan. Meticulous risk management was not done in the projects, which is understandable because EA work is a continuous process, not a project. Although EA work is, especially in the public administration sector, often considered as a project because of the funding models. The architects themselves treat EA as a process, possibly neglecting project management. It is also possible that the EA work is not supervised properly because EA projects are not considered as important as, for example, procurement projects.

This leads to our second proposition:

Proposition 2: Communication and collaboration should be addressed in the project risk management and mitigated explicitly by a communication plan and collaboration model.

Correspondingly, prior studies have identified obsolete and inadequate EA management documentation as a risk [31, 33]. Examples of risks related to the EA projects' communication and collaboration are: sufficient and varied expertise in the EA team (Project A), communication with stakeholders (Projects B and C), the architecture definition is understandable to management and developers (Projects A, B, and C), and a communication plan is missing (Projects A, B, and C). These risks can be managed by identifying sufficient resources in a project plan, designing a communication plan for the EA project, and creating dedicated architecture documents for management and developers.

It is also necessary to better prepare the stakeholders for evidently conflicting expectations. Banaeianjahromi and Smolander [7] recommended that before initiating the EA

project, increasing the personnel's trust, motivating them to collaborate, placing EA on the highest level of the organization, and ensuring that an EA team also consists non-EA experts are vital for success. The managers should also examine workflows and how the teams work [11]. These suggestions can be seen as non-technical meta-principles for EA. While Haki and Legner [19] identified some EA meta-principles, they focus on EA techniques and the quality of EA artifacts: integration, data consistency, standardization, compliance, technology independence, modularity, reusability, and usability.

This leads to our third proposition:

Proposition 3: Ensuring efficient communication and collaboration should be defined as an architecture principle in the architecture definition document. The definition should include a statement, rationale, and implications of the principle.

Contrary to Haki and Legner, we propose a communication and collaboration principle to guide architecture design and evolution [19]. Project C's architectural principles included communication and collaboration issues. Projects A and B shared their architecture principles. None did involve the communication and collaboration principle, although its necessity was acknowledged as a side note. In Project C, the management did not sufficiently consider the principle, and the architecture boards at Projects A and B did not adopt it as a principle. The TOGAF version 10, de facto EA framework, provides examples of architecture principles. Neither does it contain such a principle. As often failing EA projects demonstrate, communication and collaboration are severe problems in EA work and should thus be emphasized as an EA principle. EA projects are no different from other development projects in terms of structure or project management, so they also require proper project planning, including resourcing, risk management, and communication plans. Explicitly described the collaboration model where the stakeholders' roles and responsibilities are set, strengthens and eases project management, and mitigates communication and collaboration risks. Möhring et al. [33] argued that mature enterprise architecture management is a prerequisite for successful EA projects. One unanticipated result was that enterprise architecture management have been neglected in these projects. However, our study did not examine whether the project management was deficient.

6 Conclusion

Earlier research suggests that communication and collaboration problems must be solved to create impactful EA artifacts [6, 7, 13]. In this paper, we studied how contextual communication and collaboration problems are addressed in the EA projects.

Our projects used EA to manage their digital transformation processes. In Project A, collaboration with other stakeholders improved. In Project B, communication and collaboration problems were solved by expanding the project to cover 29 organizations. In both projects, the actions improved commitment to digital transformation. In Project C, collaboration with the responsible lawyer and the strategy group influenced the strategic goal to build a new organizational structure and an agency, which form the core of the future service ecosystem.

Our observations unveiled the consequences of the project resource reallocation and of changing the work practices. We then built three generic propositions for practitioners

to avoid the problems. Propositions 1, 2, and 3 are targeted for project management, and the third proposition is also for senior EA architects. We showed that EA practitioners have to be prepared to manage emerging communication and collaboration issues consciously and actively.

In general, enterprise architecture management is pivotal in the success of EA projects [33]. Shanks et al. [40] found that EA service capability and EA governance both have a positive impact on the success of EA projects. [2] argued for the importance of EA modeling, EA planning, EA implementing, and EA governance capabilities. However, we argue that communication and collaboration is a threshold resource in EA projects. In this respect, our three propositions concretize the argument.

We provide theoretical and practical contributions. For theory, our propositions are a starting point for future research and to study, for example, their relation to Shanks et al. [40] or Ahlemann et al. [2] capabilities. Also, our model of analysis (Fig. 5) shows some relationships with actions and their consequences. It thus provides more understanding about the EA benefit realization practices c.f. [35, 43]. For practice, the propositions provide concrete, immediately applicable advice.

This study has some limitations. First, our research method, ethnographic observations, is subjective as the first author was living the daily life of the projects. The information was extracted from the perspective of only one person, who was involved in the actions and was not only a passive observer. He influenced the data collection by selecting what to collect and record, and his memory and potential biases have probably limited what can be reviewed in the analysis phase. Although we have tried to minimize over-subjectivity and the problems of accidental misanalyses by first writing the storyline of activities and then analyzing the storyline, and by constantly reflecting on the findings among the authors, subjectivity is still there. However, as our purpose was to analyze only one problem and how it is dealt with such potential subjective bias is minimal. Second, the context, the Finnish public sector, may set some limitations. The propositions are not related to cultural or administrative issues, but they are generic and can be applied in other contexts. The third limitation is the focus on one problem type only. However, the EA problems are intertwined when they occur, and their interaction matters [7]. This relation is left for future research.

References

1. Agostino, D., Arnaboldi, M., Lema, M.D.: New development: COVID-19 as an accelerator of digital transformation in public service delivery. *Public Money Manage.* **41**(1), 69–72 (2021). <https://doi.org/10.1080/09540962.2020.1764206>
2. Ahlemann, F., Legner, C., Lux, J.: A resource-based perspective of value generation through enterprise architecture management. *Inf. Manage.* **58**(1), 103266 (2021). <https://doi.org/10.1016/j.im.2020.103266>
3. Ahlemann, F., Stettiner, E., Messerschmidt, M., Legner, C.: *Strategic Enterprise Architecture Management: Challenges, Best Practices, and Future Developments*, 1st edn. Springer-Verlag, Heidelberg (2012)
4. Ajer, A.K., Hustad, E., Vassilakopoulou, P.: Enterprise architecture operationalization and institutional pluralism: the case of the Norwegian hospital sector. *Inf. Syst. J.* **31**(4), 610–645 (2021). <https://doi.org/10.1111/isj.12324>

5. Banaeianjahromi, N.: The role of top management commitment in enterprise architecture development in governmental organizations. *Complex Syst. Inform. Model. Q.* **17**, 95–113 (2018). <https://doi.org/10.7250/csimq.2018-17.05>
6. Banaeianjahromi, N., Hekkala, R.: Factors influencing communication and collaboration in enterprise architecture development. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 6062–6071 (2019)
7. Banaeianjahromi, N., Smolander, K.: Lack of communication and collaboration in enterprise architecture development. *Inf. Syst. Front.* **21**(4), 877–908 (2019). <https://doi.org/10.1007/s10796-017-9779-6>
8. Baskerville, R.L.: Investigating information systems with action research. *Commun. Assoc. Inf. Syst.* **2**(1), 19 (1999). <https://doi.org/10.17705/1CAIS.00219>
9. Bertelsen, O. W., Bødker, S.: Activity theory. In: HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science, pp. 291–324 (2003). <https://doi.org/10.1016/B978-155860808-5/50011-3>
10. Bokolo, A., Jr.: Managing digital transformation of smart cities through enterprise architecture – a review and research agenda. *Enterp. Inf. Syst.* **15**(3), 299–331 (2021). <https://doi.org/10.1080/17517575.2020.1812006>
11. Brée, T., Karger, E.: Challenges in enterprise architecture management: overview and future research. *J. Govern. Regul.* **11**(2), 355–367 (2022). <https://doi.org/10.22495/jgrv11i2siart15>
12. Chung, S.J., Shim, O., Kim, S.: Analysing the status of communication designers' role and interdisciplinary collaboration. *Des. J.* **25**(3), 397–415 (2022). <https://doi.org/10.1080/14606925.2022.2058682>
13. Dang, D.D., Pekkola, S.: Problems of enterprise architecture adoption in the public sector: root causes and some solutions. In: Rusu, L., Viscusi, G. (eds.) *Information Technology Governance in Public Organizations*. ISIS, vol. 38, pp. 177–198. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58978-7_8
14. Datta, P., Walker, L., Amarilli, F.: Digital transformation: learning from Italy's public administration. *J. Inf. Technol. Teach. Cases* **10**(2), 54–71 (2020). <https://doi.org/10.1177/2043886920910437>
15. Gabryelczyk, R.: Has COVID-19 accelerated digital transformation? Initial lessons learned for public administrations. *Inf. Syst. Manage.* **37**(4), 303–309 (2020). <https://doi.org/10.1080/10580530.2020.1820633>
16. Goerzig, D., Bauernhansl, T.: Enterprise architectures for the digital transformation in small and medium-sized enterprises. *Procedia CIRP* **67**, 540–545 (2018). <https://doi.org/10.1016/j.procir.2017.12.257>
17. Gong, Y., Janssen, M.: The value of and myths about enterprise architecture. *Int. J. Inf. Manage.* **46**, 1–9 (2019). <https://doi.org/10.1016/j.ijinfomgt.2018.11.006>
18. Gong, Y., Janssen, M.: Why organizations fail in implementing enterprise architecture initiatives? *Inf. Syst. Front.* **25**(4), 1401–1419 (2022). <https://doi.org/10.1007/s10796-022-10298-x>
19. Haki, K., Legner, C.: The mechanics of enterprise architecture principles. *J. Assoc. Inf. Syst.* **22**(5), 1334–1375 (2021). <https://doi.org/10.17705/1jais.00696>
20. Jusuf, M. B., Kurnia, S.: Understanding the benefits and success factors of enterprise architecture. In: HICSS 2017, pp. 4887–4896 (2017). <https://doi.org/10.24251/HICSS.2017.593>
21. Kaisler, S.H., Armour, F., Valivullah, M.: Enterprise architecting: critical problems. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, pp. 224b–224b (2005). <https://doi.org/10.1109/HICSS.2005.241>
22. Kaisler, S.H., Armour, F., Sprague, R., Bui, T.X.: 15 years of enterprise architecting at HICSS: revisiting the critical problems. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 4807–4816 (2017)

23. Korhonen J.J., Halén, M: Enterprise architecture for digital transformation. In: IEEE Conference on Business Informatics, vol. 1, pp. 349–358 (2017). <https://doi.org/10.1109/CBI.2017.45>
24. Kotusev, S., Kurnia, S., Dilnutt, R.: Enterprise architecture artefacts as instruments for knowledge management: a theoretical interpretation. *Knowl. Manag. Res. Pract.* **21**(3), 594–606 (2023). <https://doi.org/10.1080/14778238.2021.1999183>
25. Kotusev, S., Kurnia, S.: The theoretical basis of enterprise architecture: a critical review and taxonomy of relevant theories. *J. Inf. Technol.* **36**(3), 275–315 (2021). <https://doi.org/10.1177/0268396220977873>
26. Kotusev, S.: Enterprise architecture and enterprise architecture artifacts: questioning the old concept in light of new findings. *J. Inf. Technol.* **34**(2), 102–128 (2019). <https://doi.org/10.1177/0268396218816273>
27. Kurnia, S., Kotusev, S., Shanks, G., Dilnutt, R., Milton, S.: Stakeholder engagement in enterprise architecture practice: what inhibitors are there? *Inf. Softw. Technol.* **134**, 106536 (2021). <https://doi.org/10.1016/j.infsof.2021.106536>
28. Lamanna, F., Kurnia, S.: Digital transformation of superannuation and the role of enterprise architecture. In: PACIS 2022 Proceedings, p. 214 (2022)
29. Lemmetti, J., Pekkola, S.: Understanding enterprise architecture: perceptions by the Finnish public sector. *Int. Conf. Electron. Govern.* **7443**, 162–173 (2012). https://doi.org/10.1007/978-3-642-33489-4_14
30. Levina, A.I., Borremans, A.D., Lepekhin, A.A., Kalyazina, S.E., Schröder, K.M.: The evolution of enterprise architecture in scopes of digital transformation. *IOP Conf. Ser. Mater. Sci. Eng.* **940**(1), 012019 (2020). <https://doi.org/10.1088/1757-899X/940/1/012019>
31. Löhe, J., Legner, C.: Overcoming implementation challenges in enterprise architecture management: a design theory for architecture-driven IT management (ADRIIMA). *ISEB* **12**(1), 101–137 (2014). <https://doi.org/10.1007/s10257-012-0211-y>
32. Mirsalari, S.R., Ranjbarfard, M.: A model for evaluation of enterprise architecture quality. *Eval. Program Plann.* **83**, 101853 (2020). <https://doi.org/10.1016/j.evalprogplan.2020.101853>
33. Möhring, M., Keller, B., Schmidt, R., Sandkuhl, K., Zimmermann, A.: Digitalization and enterprise architecture management: a perspective on benefits and challenges. *SN Bus. Econ.* **3**(2), 46 (2023). <https://doi.org/10.1007/s43546-023-00426-3>
34. Niemi, E., Pekkola, S.: Using enterprise architecture artefacts in an organization. *Enterp. Inf. Syst.* **11**(3), 313–338 (2017). <https://doi.org/10.1080/17517575.2015.1048831>
35. Niemi, E., Pekkola, S.: The benefits of enterprise architecture in organizational transformation. *Bus. Inf. Syst. Eng.* **62**(6), 585–597 (2020). <https://doi.org/10.1007/s12599-019-00605-3>
36. Randall, D., Harper, R., Rouncefield, M.: *Fieldwork for Design: Theory and Practice*. Springer Verlag, London (2007). <https://doi.org/10.1007/978-1-84628-768-8>
37. Rouvari, A., Pekkola, S.: Contextuality and temporality of enterprise architecture problems: a comparative case study. In: Sales, T.P., et al. (eds) *Enterprise Design, Operations, and Computing*. LNBIP, vol. 466, pp. 180–197. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-26886-1_11
38. Scupola, A., Mergel, I.: Co-production in digital transformation of public administration and public value creation: the case of Denmark. *Gov. Inf. Q.* **39**(1), 101650 (2022). <https://doi.org/10.1016/j.giq.2021.101650>
39. Seppänen, V.: *From problems to critical success factors of enterprise architecture adoption*. University of Jyväskylä, Jyväskylä (2014)
40. Shanks, G., Gloet, M., AsadiSomeh, I., Frampton, K., Tamm, T.: Achieving benefits with enterprise architecture. *J. Strateg. Inf. Syst.* **27**(2), 139–156 (2018). <https://doi.org/10.1016/j.jsis.2018.03.001>

41. Simon, D., Fischbach, K., Schoder, D.: Enterprise architecture management and its role in corporate strategic management. *IseB* **12**(1), 5–42 (2014). <https://doi.org/10.1007/s10257-013-0213-4>
42. Soto-Acosta, P.: COVID-19 pandemic: Shifting digital transformation to a high-speed gear. *Inf. Syst. Manag.* **37**(4), 260–266 (2020). <https://doi.org/10.1080/10580530.2020.1814461>
43. Tamm, T., Seddon, P.B., Shanks, G.: How enterprise architecture leads to organisational benefits. *Int. J. Inf. Manag.* **67**, 102554 (2022). <https://doi.org/10.1016/j.ijinfomgt.2022.102554>
44. Tamm, T., Seddon, P.B., Shanks, G., Reynolds, P.: How does enterprise architecture add value to organisations? *Commun. Assoc. Inf. Syst.* **28**, 10 (2011). <https://doi.org/10.17705/1CAIS.02810>
45. van de Wetering, R., Kurnia, S., Kotusev, S.: The effect of enterprise architecture deployment practices on organizational benefits: a dynamic capability perspective. *Sustainability* **12**(21), 8902 (2020). <https://doi.org/10.3390/su12218902>
46. Ylinen, M., Pekkola, S.: Jack-of-all-trades torn apart: Skills and competences of an enterprise architect. In: Proceedings of the 28th European Conference on Information Systems (ECIS). Research paper 13 (2020)
47. Yin, R.K.: *Case Study Research: Design and Methods*, vol. 5. Sage, California (2009)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Navigating ICT In-House Procurement in Finland: Evaluating Legal Frameworks and Practical Challenges

Reetta Ghezzi^(✉), Minnamaria Korhonen, Hannu Vilpponen,
and Tommi Mikkonen

University of Jyväskylä, Jyväskylä, Finland
{reetta.k.ghezzi,hannu.v.vilpponen,tommi.j.mikkonen}@jyu.fi

Abstract. In-house procurement is a controversial issue in the field of public procurement. Simply put, such procurement allows overlooking certain aspects of fair and equal treatment of vendors. This paper presents qualitative research on in-house ICT procurement within Finnish municipalities. Semi-structured interviews were conducted to gather insights from municipal stakeholders. Using grounded theory approach, data analysis shows intricate dynamics between Finnish municipalities and in-house entities associated with them. Still, it is clear that the legal framework governing in-house procurement remains intricate and debated.

Keywords: Public procurement · In-house companies · Software acquisition · Public sector information systems

1 Introduction

The public sector is a large consumer of ICT systems and services [3]. For example, the Finnish government alone made ICT purchases worth over EUR 1000 million in 2022 [2]. In addition, Finnish municipalities, joint municipal authorities, and parishes made ICT purchases worth almost EUR 1500 million [2]. With this in mind, the Public Procurement Directive [9] encourages EU Member States to adopt transparent and pro-competitive procurement practices. Public bodies may adopt vast procurement opportunities to achieve these requirements. The first option is to tender the purchase publicly [14]. The second option involves in-house procurement or procurement from other stakeholder units, which falls outside the scope of public procurement law, in this case [14].

So-called in-house companies are owned by public organizations. Their role in public sector procurement has recently attracted much attention, as transparency and openness in in-house procurement can be difficult to implement [12]. Moreover, in-house procurement can lead to difficulties in obtaining information and data from in-house companies. Finally, legal interpretations of in-house status can be unclear [12].

In this paper, we study how much Finnish municipalities rely on in-house procurement and why municipalities do or do not use in-house procurement. Sixteen semi-structured interviews with procurement and ICT key persons in municipalities were used to collect the research data. The interviews were conducted face-to-face or by video conference, whichever was most convenient for the interviewee. The paper is structured as follows. Section 2 presents the background of this work. Section 3 introduces the research setup, and Sect. 4 lists the key findings. Section 5 discusses the key findings. Section 6 draws some final conclusions.

2 Background and Motivation

The Public Procurement Act [9] governs public acquisitions. However, it does not apply when a contracting authority, for example, a municipality, makes a procurement from a company it owns, called an in-house company, provided that the in-house company is formally separate for policy-making purposes, has a controlling interest by the municipality and conducts only a limited amount of business with external parties [1]. Procurement Directive allows 20 percent of turnover to go outside the owners of the in-house company [9]. However, in Finnish law, the threshold for outselling is stricter. Public Procurement Act specifies that 5 percent and EUR 500,000 limits for outselling apply based on the in-house entity's turnover three years before the agreement [1]. However, these limits don't apply when there's no market-based operation to execute the services. Whether the market-based operations exist is determined by the responses to a transparency declaration [1].

Procurement units that own the in-house company must have decisive authority in the in-house company [9]. The Public Procurement Act defines joint-decisive authority as when all contracting entities have representatives in the in-house company's executive organs and collectively make strategic decisions, with the condition that the in-house company operates in the interests of the controlling contracting entities [18]. In addition, the Public Procurement Act states that it does not apply when an in-house company is a procurement unit itself and procures goods or services from another procurement unit, which exercises controlling interest in the in-house company or another entity under the same controlling interest [1]. This option is the so-called in-house sisters' arrangement in Finland. The recent judgment of the EU Court of Justice (ECJ) in the *Sambre & Biesme* case [23] would seem to contradict the article in the Finnish Public Procurement Act or at least guide how to interpret Section 15 of the Procurement Act. In this case, the need for real representation in the in-house company's board or decision-making bodies was emphasized, possibly contradictory to the Procurement Act. Ownership of the shares alone did not guarantee decisive authority in the in-house sister arrangement, even if the other procurement unit had decisive authority in the in-house company. This shows that factors related to the in-house company's governance and joint-decisive authority can significantly impact assessing its in-house status.

Some other ECJ judgments depict how to evaluate adequate in-house positions. In the *Parking Brixen* case [22], the municipality lacked sufficient decisive authority in the in-house company, rendering the company not part of the municipal in-house. Similarly relating to the evaluation of the owner's sufficient decisive authority, the *Carbotermo and Concorzia Alise* case [6] considered how the strong dominant position of majority shareholder affects the legal position of other shareholders in an in-house company. The risk of conflict of interest is high, and it can influence the in-house company's legal position. If only one or a few shareholders have real decisive authority, the objectives of the other owners are not given space; their realization is uncertain and, therefore, it may create a situation where those with little or no decisive authority do not have a real in-house position in the company they own.

Recent public discussion has been raised over the in-house position as habitual practice through ownership and a somewhat fictitious demonstration of decisive authority. Within similar themes, in *Econord's* case, the significance of structural and operational control in assessing in-house status was highlighted [10]. Formal ownership is insufficient to ensure in-house status [10]. This suggests that even small shareholders should have sufficient joint-decisive authority over the in-house company's operations, and in-house position cannot be presented merely on paper. For example, the largest Finnish in-house company, *Kuntien Tiera*, has 398 owners. As methods of decisive authority, *Kuntien Tiera* states that the owners steer *Kuntien Tiera's* activities in the general assembly and the board of directors, as well as the developing *Kuntien Tiera's* service offerings in six different steering groups [21].

Based on these legal cases, it is evident that the importance of real decisive authority and ownership in the in-house company is significant. In addition to ownership share, importance is also given to control, structure, decision-making, and genuine representation in the in-house company's operations. It is important to assess these factors as a whole when evaluating the legal status of an in-house company.

The in-house arrangement can be challenging to interpret for municipalities [12]. Despite clear guidelines provided by case law, there is a significant variation in their interpretation in practice [12]. The legal setup surrounding in-house procurement is a critical issue discussed in the literature. In Poland, where stricter in-house procurement criteria have been implemented, the debate is polarised between supporters and opponents [15]. Opponents seem to question whether in-house practice aligns with the goals set in legislation [15]. Similarly, *Burgi and Koch* [5] evaluate the Public Procurement Directive article 11 and suggest that lowering the criteria for in-house procurement could be a way to prevent legal mismatch and confusion in the field.

In practical applications, in-house procurement may benefit smaller municipalities by reducing the bureaucracy involved in contracting and contract implementation costs [20]. However, it has been questioned whether the upcoming, now-current directives will create a procurement market that does not have to obey and is not controlled by procurement norms [16]. The concerns are that the

upcoming directives will exclude private service providers from the competition if the in-house exception is accepted [16]. Similar concerns have been raised in Finland as well. The Confederation of Finnish Industries has raised concerns that the current in-house practice distorts the market and has taken steps to address these concerns through a request for measures to the practices from the Competition and Consumer Authority [11]. Baciú suggests that public bodies should not be able to avoid transparent procedures and contract directly with other public bodies, except in rare and limited situations to preserve fair competition [4]. The Confederation of Finnish Industries and the Finnish Competition and Consumer Authority also take the same view in their proposals [11, 17].

The literature concludes the current procurement directive inhibits opening up the national procurement markets and fosters direct awarding in public contracts, even if the underlying purpose is the opposite. The challenges surrounding in-house procurement for public entities highlight the need for continued examination and clarification of legal frameworks and in-house procurement criteria.

3 Research Approach

The research will focus on municipalities and well-being services counties in Finland. The research questions for this study are:

- RQ1: When should a public organization procure from in-house and when to procure from the market?
- RQ2: How much real decisive authority do public sector organizations and municipalities hold in the in-house arrangement?

Data Collection. The primary data collection method was semi-structured interviews with sixteen key stakeholders from municipalities and well-being services counties. The interviews were conducted face-to-face or via video conferencing. The approach to design the interviews was constructive [7], and therefore, the interviews were recorded because the aim was to preserve the details such as participant’s tempo and tone as precisely as possible. However, only the audio of all interviews was recorded, and otherwise, for observation purposes, notes taken during the interview were relied upon. According to Glaser, the notes capture what is needed without losing the detail [13]. During the analysis phase of this study, it was found that the recordings were an excellent supplement for interpreting the interviewee’s attitudes and assumptions of in-house procurement. Especially when discussing more difficult topics, such as the legal status of in-house companies or the role of the small owner, the recordings helped to understand the hesitation and uncertainty. Only one interviewee requested that the video link not be used. Transcribed interview data was loaded into the *atlas.ti* software for coding.

All participants were professionals in their field, either in public procurement in general, ICT procurement and its management, or in the financial management of the organization. All participants were involved in in-house procurement in one way or another (Table 1).

Table 1. Interview participants.

Participant	Abbreviation	Position	Field	Minutes
Participant 1	P1	Chief Financial Officer	Administration	107
Participant 2	P2	Procurement Manager	Procurement / ICT	49
Participant 3	P3	City Director	Administration	53
Participant 4	P4	Head of Procurement Expert Group	Procurement / ICT	49
Participant 5	P5	Chief Digital Officer	ICT	67
Participant 6	P6	City Auditor	Administration	57
Participant 7	P7	Division Director	ICT	51
Participant 8	P8	Procurement Specialist	Procurement / ICT	56
Participant 9	P9	Procurement Manager	Procurement	56
Participant 10	P10	Support Services Director	ICT	75
Participant 11	P11	Chief Information Officer (CIO)	ICT	53
Participant 12	P12	Chief Information Officer (CIO)	ICT	61
Participant 13	P13	Welfare County Director	Administration	57
Participant 14	P14	Municipal Director	Administration	105
Participant 15	P15	Administrative Director	Administration	105
Participant 16	P16	Chief Financial Officer	Administration	105

Analysis. The grounded theory (GT) approach suits topics lacking relevant research or where a new perspective is desired [26]. The practice of ICT in-house procurement is an unexplored area in Finland, except for the request for measures [17] and report [24] by the Consumer and Competition Authority and surveys conducted by Confederation of Finnish industries [19]. Fresh European in-house procurement research is also extremely limited.

The GT approach to research involves systematically coding and classifying data [25]. GT stands apart from other qualitative research methods primarily in its approach to analysis, while data collection methods can vary. Typically, GT involves constructing theories based on interview data, with data collection continuing until saturation is reached [26]. Saturation means no new information relevant to the developing theory is emerging [8].

In this research, the coding followed a constructive approach to the grounded theory [7]. The open coding stage included initial coding and sometimes codes that emerged from the participants' narratives, known as "in vivo" coding. This constituted the first analysis phase, establishing a data-driven initial sorting [7]. The initial codes facilitated comprehension of the interview material and the intended meanings conveyed by the interviewees. Subsequently, after each interview, a comprehensive review of the material and codes was conducted to verify that the codes consistently conveyed the same concept across all interviews. Charmaz underscores the significance of constant comparison within GT, a methodology involving the comparison of categorized data instances within the same category [7]. As outlined by Urquhart in 2023, this approach aims to assess the compatibility and efficacy of the identified categories [26].

As coding progressed in the study, focused coding advanced the analysis to a more theoretical direction with conceptualization, for example, recognizing where the initial codes lead the process:

“Feels disempowered in cooperation.” – “Signs of insufficient decisive authority.”

After focused coding, thoughts arose about the relationships between these codes. These relationships were marked utilizing the *atlas.ti* memo and grouping function. At this point, the axial coding stage [7] and the selective coding stage were somewhat parallel processes [26] [7]. The phase of seeking common themes and grouping categories helped us understand the causation relationships.

The significance of theoretical notes in understanding relationships was emphasized and aided in forming an overall picture. Coding, categorization, and grouping were flexible throughout the analysis, and changes occurred until the key categories were fully saturated and no new codes emerged. Ultimately, 996 quotations were selected from the material and categorized under 149 codes. It should be noted that around 700 additional quotations were coded related to clusters, such as themes concerning the organization of public entities in procurement, monitoring, and measurement of procurement, ICT project management, public organization management, and system solution-related themes.

4 Results

4.1 Reasons for ICT In-House Procurement

There are several characteristics by which in-house procurement can be justified. It allows sharing of the risk and costs of producing certain widely used services, as well as due to different financial capacities of public sector organizations. Below, we present the key reasons for using ICT in-house procurement found in this study.

ICT in-house Companies are Widely Utilized due to Shortcomings in the Existing Market. Sometimes, only a few (and sometimes no) bids are received for ICT procurement. Then, in-house companies are the sole providers capable of offering support services to public sector organizations, such as systems for managing human resources and payroll. Municipalities and welfare service counties believe it would be a welcome addition if market players extended their services to the sector where ICT in-house companies currently operate. Available solutions and service production encounter challenges believed to be alleviated through increased competition within the sector, thereby providing alternative solutions to meet various needs.

In addition, interviews reveal that ICT in-house companies are extensively utilized for ICT hardware and equipment procurement, even though this type of procurement is typically considered straightforward. Some public organizations procure equipment through in-house channels, driven by the belief that the market cannot provide the necessary volumes. However, certain public organizations have realized that ICT equipment obtained through in-house procurement tends to be more expensive than market-based solutions. These organizations emphasize that entities should explore what markets can offer to ensure the most responsible use of public funds.

ICT in-house Procurement is Faster than Competitive Bidding.

Obtaining products and services from an ICT in-house company is a straightforward process. Local government sectors often have limited resources to engage in bidding, typically alongside employees' regular duties, often in collaboration with the procurement team or center. However, expertise must come from within the specific sector to oversee the bidding process.

ICT in-house procurement can enhance municipal operations by agilely utilizing resources, time, and expertise required for daily operations when the cooperation is optimal. Compared to competitive bidding, ICT in-house procurement is swift and convenient for municipalities, especially for fulfilling simple needs. Interviews also underscore that ICT in-house procurement is considered a reliable method, particularly in smaller organizations where the likelihood of legal disputes is reduced. Competitive bidding is considered burdensome and error-prone, making ICT in-house procurement a suitable option, particularly when resource constraints are a factor.

Finally, ICT in-house procurement played a pivotal role in the recent establishment of well-being services in counties instead of municipalities, which had organized the services previously. The timeline was so strict that would have been impossible to tender market-based competitive bidding, as per procurement law. Furthermore, central procurement organizations lacked the capacity for proper competitive bidding while establishing well-being services in counties was under construction. Then, through ICT in-house companies, well-being services in counties were operationalized within a tight 1.5-year timeframe.

Resources and Expertise Within Public Organizations may often Prove Inadequate.

More than half of the interview participants believe that public organizations lack personnel who understand the ICT needs of the sectors well enough to support the creation of coherent system configurations. Additionally, these organizations often lack personnel who can simultaneously grasp the diverse requirements of competitive bidding in accordance with procurement laws. When a public organization lacks both ICT and procurement expertise, ICT in-house procurement becomes a viable option for acquiring products and services simply because everything seems to be readily available off the shelf.

The Desire is to Centralize Collaboration in One Location and Obtain Shared and Standardized ICT Systems Through in-house Procurement.

Local governments and well-being services counties believe that certain needs within public organizations are quite similar, particularly those related to support services. Municipalities seek to harness the benefits of collaboration and shared systems to achieve cost-efficiency and agility in such cases. Furthermore, system compatibility among municipalities facilitates rapid service delivery and error correction. The ICT in-house practice may not always meet this need, leading some municipalities to purchase the same system offered by ICT in-house directly from the system provider in an attempt to resolve issues directly with the supplier.

ICT in-house Procurement is Needed to Enhance Collaboration. ICT in-house companies have emerged because clear, distinguishable functions within

Table 2. Key issues in in-house procurement

Key issues in in-house procurement in this research.	Percentage of how many interviewees hold the opinion.
No real decisive authority in in-house company	81 %
Small shareholder: Small buyer	75 %
In-house companies are currently too large	69 %
Poor reputation	63 %
Expensive solutions	63 %
In-house ownership, shareholder position	63 %
Insufficient expertise in the system development and/or procurement	56 %
Contracts with in-house companies are not binding or they do not exist	44 %
Exiting from in-house is demanding	44 %
In-house: not functioning as it should	44 %
Chain of command doesn't work	38 %
Service and system development are slow	31 %
Poor quality of relationships	31 %
Trust has been eroded	25 %
Service does not meet the agreed terms	25 %

public organizations are identified for collaborative production with other entities with similar needs. An example of such a function could be payroll processing. The goal is to enhance the efficiency of public organizations by centralizing and sharing production costs with other stakeholders while freeing up internal resources. Additionally, centralization aims to harness expertise-related benefits, allowing for the incorporation of necessary expertise from external sources, where such expertise is perceived to be concentrated within that specific function. The ICT in-house practice also aims to ensure the security of critical system operations and their continuous functionality.

4.2 Key Problems Related to ICT In-House Companies

Despite the benefits, some problems arise in the context of ICT in-house companies. Table 2 provides an overview of key issues related to ICT in-house companies. In summary, insufficient decisive authority, the position of minority shareholders, the rapid expansion of ICT in-house companies, damaged reputation, costly solutions, deficiencies in contract practices, and issues related to ownership shares emerge as central problems based on the study. This section discusses the challenges within ICT in-house companies and their potential sources.

Challenges Related to Insufficient Decision-Making Authority and the Legal Position of Small Shareholders. In municipalities and well-being services counties, there is a comprehensive understanding of how an in-house position could be achieved through procurement law. Ownership in the in-house company and decisive authority are central for the evaluation, as shown in Fig. 1. All organizations in this research are small shareholders in the central in-house companies which we took for reference. Wide consensus exists about marginal ownership, seen as an established practice, and interviewees believe there is hardly room to interpret the matter differently.

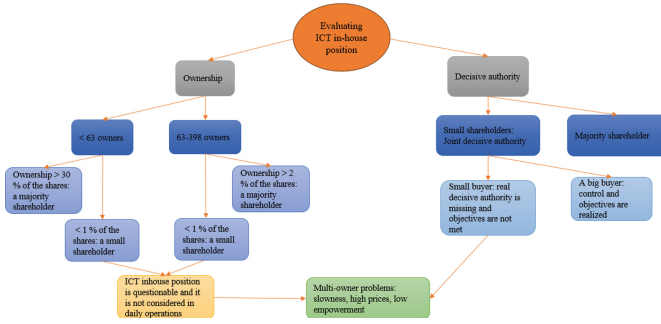


Fig. 1. Evaluation of the in-house position in studied organizations.

The problem arises from the unclear interpretation of sufficient decisive authority, which is also evident in interviews through varying interpretations. Within the interviews, three interpretations existed, as presented in Fig. 2. Joint-decisive authority divides opinions. Most interviewees depict that mechanisms work with even a small ownership stake or nominal authority, and a small ownership stake is deemed sufficient for the in-house position. The difference arises when considering the purchase sizes mentioned by interviewees. Large buyers feel that authority works and collaboration with ICT in-house companies is immediate. Problems are reacted swiftly, and organizational goals are achieved through in-house ICT collaboration. Some large buyers actively participate in decision-making bodies. One large buyer expressed thoughts about ownership not guaranteeing sufficient decisive authority:

“To me, these shares and decisive authorities and such; the idea that ownership gives you a certain position, I might not fully buy it. And then I think, are these matters as extensive as they have been portrayed in public.” (P3)

Some large buyers do not directly engage in the decision-making of ICT in-house companies, but they trust that shared authority is sufficient for evaluating the in-house position:

“Well, there’s a well-established legal practice in Finland that you don’t need to think about; if you have an in-house service provider and you’ve delved into it a bit, then you don’t need a separate evaluation. Well-established legal practice means that there’s such an in-house service provider where the owners exercise decisive authority together. The legislation is quite clear. It doesn’t require any extraordinary evaluation. Of course, if the Competition and Consumer Authority asks, then we hire a lawyer who writes 10 pages about how it (joint-decisive authority) is done, but the matter is just this simple.” (P1)

All small shareholders with significant purchases consider ICT in-house operations to align with their goals and find their authority in in-house companies effective. This is why the situation becomes problematic when we consider

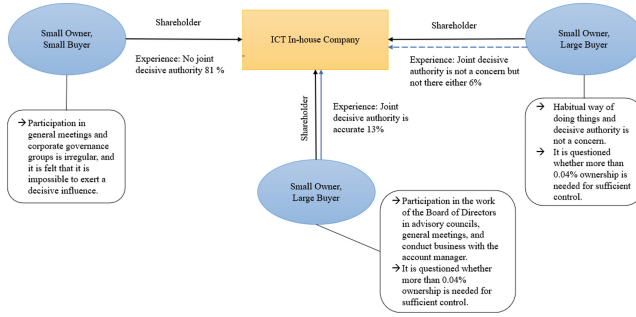


Fig. 2. Recognized differences between minority shareholders’ views about decisive authority and ownership.

the experiences of small owners with small purchases. The views of large and small buyers are conflicting, as small buyers perceive there to be no real decisive authority in the ICT in-house companies:

“Almost non-existent (decisive authority mechanisms). We own 0.01 % shares there, and then we’re supposed to have decisive authority. If this counts as an in-house company as per procurement law, I’ve also thought a lot about how this can be.” (P4)

Again, the in-house position is evaluated based on ownership and decisive authority, yet the small buyer’s experience differs significantly from that of larger buyers. Consistently, small buyers question whether they possess a sufficient number of shares to attain proper decision-making authority within the in-house company, here we see how these two factors are assessed as equivalent criteria in determining the position of in-house companies, which differs from the reports of large buyers.

“Well, the influence there is really small, that they are owner-managed companies, but each owner has such a small share that we don’t know who actually controls it.” (P5)

In addition, small buyers have refrained from participating in situations where joint decision-making authority could be demonstrated because it has been deemed futile:

“None of us have actually attended the general meetings anymore. Formally speaking, there are these owner meetings where strategic matters are discussed, where all over 100 shareholders use their weighty vote, and there’s also a formal board member representing minority shareholders. I don’t really feel that we have concrete influence over it (in-house company).” (P6)

In summary, the majority of small shareholders with modest purchases believe that they lack significant authority over ICT in-house companies. Moreover, all study participants view ICT in-house companies as part of the market since the control mechanism does not function as intended for their own units. If the same objectives were applied to ICT in-house companies as for their own units, they could be considered an integral part of their own production.

Fast Expansion of the ICT In-House Companies. The interview responses suggest a significant increase in the number of owners of ICT in-house companies in recent years, largely due to mergers of smaller regional entities into larger national ones. This growth, particularly in the context of the central ICT in-house companies examined in the study, has been substantial, especially regarding the number of minority shareholders. The interviews also shed light on the challenges minority shareholders face, particularly those with smaller purchases, compared to majority shareholders. Notably, municipalities have observed that larger cities with greater ownership and purchasing power tend to receive priority in terms of the systems offered and their quality. This bias towards major owners often results in the goals of minority shareholders with limited influence within the in-house company not being met. As a consequence, the existence of multiple owners poses considerable challenges in achieving common objectives. In the central ICT in-house companies studied, as well as those discussed in the interviews, the ownership structure varies widely, ranging from 47 to 398 owners. It is noteworthy that all participating organizations hold a minority ownership position in these ICT in-house companies, with ownership stakes spanning from 0.00 to 1.00 percent of the shares.

Significant Variations in ICT In-House Companies' and Owner's Contract Practices. The study highlights significant variations in contract practices between ICT in-house companies and their owners. During the establishment of well-being services counties, some municipalities lacked contracts with ICT in-house companies, posing challenges when attempting to transfer contracts to the well-being services counties. Respondents also mentioned that the most significant problems with ICT in-house companies occur when contracts are entirely absent. Addressing errors becomes nearly impossible when the party supplying the system or service is not obligated to act. In addition, uncertainties exist in contract clauses related to service levels, lacking specific obligations outlined for the owner and the ICT in-house company. While most contracts state that problem situations should be resolved through collaboration, detailed service-level descriptions with obligations typical of the private sector seem to be entirely absent. Some ICT in-house companies prefer a standardized platform for all owner contracts that all owners can access, while others draft contracts only upon request.

5 Discussion

Root Causes for Problems. Insufficient control by owners and the rapid expansion of ICT in-house companies are strongly interrelated. According to

Table 3. Antecedents, Field Experiences, and Consequences.

Antecedents A1-A4	Experiences E1-E8	Consequences C1-C5
A1. Fast expansion of ICT in-house companies	E1. Lack of decision-making power by the owners	C1. Common objectives are not met
	E2. Small owners and small buyers have a weak position	
	E3. In-house status sometimes questionable	
A2. Shortcomings in contractual practice.	E4. Service and system development is slow, reacting to issues and errors is slow.	C2 Current practice does not hold ICT in-house companies liable for errors.
	E5. Vendor lock-in with in-house company and supplier.	
	E6. Changes are almost impossible	
A3. ICT in-house companies dominate their market sector	E7. Expensive solutions.	C3. Operations are interrupted or significantly impeded.
A4. The market is not working / No competition	E8. Vendor lock-in with in-house company and supplier.	C4. High costs
		C5. Updates and changes are mandatory.

the study, there is an imbalance in the position of small shareholders, leading to problems associated with multi-ownership, such as the fact that small shareholders may not necessarily pursue common objectives. Small shareholders also hold very small ownership stakes, which raises the question of whether achieving dominant control in an ICT in-house company is structurally possible. If the interpretation is strict, the subsidiary status of ICT in-house companies might be problematic and contrary to the objectives of procurement law Sect. 15 [1].

Contractual practices vary a lot among in-house ICT companies and owners. Some ICT in-house companies have transparent contractual practices, while others have significant gaps in their contractual practices, leading to slow development of services and systems, difficulty in reacting to errors, and contracts lacking clear responsibilities for the in-house companies. ICT in-house companies dominate their market, and direct public competition rarely attracts many bids. The study indicates that 63 percent of the respondents consider solutions through in-house ICT companies expensive. However, municipalities and well-being services counties might not have any alternative but to continue with ICT in-house services, as migration costs would be too high. The lack of competition often results in price increases and decreased quality. Smaller owners are also forced to implement system updates and changes, which is relatively more expensive than larger buyers. Table 3 presents the recognized interrelationships.

Identified Preconditions for Success. When functioning properly, ICT in-house companies could bring efficiency, free up resources, and provide the necessary expertise to their owner organizations, similar to Miemec stated [20]. A prerequisite for this is that ICT in-house companies should be manageable,

ensuring the necessary structural and operational control as mandated by the law, enabling effective control of their operations. This implies that in-house ICT companies should have fewer owners yet enough to achieve economies of scale. The current Finnish government has recommended that ownership shares in in-house companies comprise a minimum of 10 percent. This proposal elicits apprehension regarding its possible detrimental impact on the well-established in-house model in Finland. More precisely, it can potentially disrupt the current in-house structure, possibly encouraging the emergence of smaller, fragmented entities with duplicated responsibilities and management functions. Importantly, this may not necessarily foster the standardization of ICT systems and services.

One interesting option has not been studied. In the Sambre & Biesme case, an in-house entity had different groups of owners with different decisive authority [23]. In the Finnish Limited Liability Companies Act, the option to allocate decisive authority differently than *one share – one vote* principle is available as well [18]. In this research, we recognized different buyer characteristics and how joint-decisive authority divides them. The shares in in-house companies are now allocated according to the population base served by the owner organization, or in the cases of well-being services counties, we did not find the justification. The purchaser groups, whether the buyer is small or large, could help to even out or create new mechanisms for how the decision-making should happen in the in-house company. This suggestion, however, needs more research to see whether it could be a viable option in practice.

Recommendations. This study identifies practices that could enhance current in-house practices and improve public sector organizations' and market actors' influence over the operations of ICT in-house companies. In the literature [5], it has been suggested that criteria for in-house procurement should be relaxed to avoid legal incompatibility and confusion. However, this study proposes a different approach since there is a lack of oversight and competition, resulting in significant national economic problems. The study reveals that most respondents perceive control over ICT in-house companies as weak, leading to slow development of services and systems, high costs, and challenges in correcting errors. The results suggest that, in certain situations, problems related to delivery can be avoided. In situations where ICT in-house companies are under the immediate control of their owners and control is closely aligned with the owners' goals, ICT in-house companies can serve as a resource to free up procurement competition. Close ownership relationships require sufficient ownership and less than fifty owners, enabling genuine structural and operation control. As a result, the procurement law needs clarification on what constitutes sufficient ownership in an in-house company. Contrary to [5], our results indicate that clear control mechanisms, strong control, and evidence of in-house status from procurement law could help reduce legal incompatibility and confusion in in-house procurement.

Threats to Validity. While GT is considered data-driven, it is impossible to completely eliminate the influence of the researcher's prior experiences and theoretical frameworks. These factors inevitably shape the analysis. Moreover,

for research to be meaningful, it should connect to previous studies and ongoing scientific discussions. Instead of strictly adhering to inductive reasoning, this research incorporates abduction (e.g. Table 3) and relies on GT theory-building characteristics. This acknowledges the role of the researcher's thinking while recognizing the importance of existing theoretical tools and context.

6 Conclusions

In conclusion, in-house procurement remains a controversial issue in public procurement. While some argue that it provides flexibility and cost savings for public authorities, others express concern about potential abuses of the exemption and the impact on fair competition. As reflected, the legal framework surrounding in-house procurement is complex and subject to ongoing debate.

This paper identified various key reasons for ICT in-house procurement and why it is important for its owners. Key problems were highlighted, and recommendations were formulated based on literature and research on practically improving operations. The research revealed valuable insights into the complex relationships between Finnish municipalities and their in-house companies. The study also touched upon the legal framework related to ICT in-house procurement, a pivotal issue in scholarly literature, emphasizing the ongoing need to review legal frameworks and in-house procurement criteria to address challenges posed to municipalities by in-house procurement.

References

1. Act on Public Procurement and Concession Contracts 1397/2016 (2016). <https://www.finlex.fi/fi/laki/kaannokset/2016/en20161397.pdf>
2. Tutki hallintoa: (in Finnish), June 2023. <https://www.tutkihallintoa.fi/julkiset-hankinnat/hankintojen-arvo/>
3. Arrowsmith, S., Hartley, K.: Public Procurement. Edward Elgar Publishing, Cheltenham (2002)
4. Baciu, I., Dragoş, D.C.: Horizontal in-house transactions vs vertical in-house transactions and public-public cooperation. *Eur. Procure. Public Priv. Partnersh. Law Rev.* **10**(4), 254–272 (2015)
5. Burgi, M., Koch, F.: In-house procurement and horizontal cooperation between public authorities. *Eur. Procure. Public Priv. Partnersh. Law Rev.* **7**(2), 86–93 (2012)
6. Carbotermo and Consorzio Alisei C-340/04. Carbotermo and Consorzio Alisei Parking Brixen GmbH v Gemeinde Brixen and Stadtwerke Brixen AG. ECJ, May 2006
7. Charmaz, K.: Constructing Grounded Theory, 2nd edn. Sage, Los Angeles (2014)
8. Dey, I.: Grounded theory. In: *Qualitative Research Practice*, pp. 80–93 (2004)
9. Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02014L0024-20180101>
10. Econord SpA v Comune di Cagno and Others Joined Cases C-182/11 and C-183/11. ECJ, November 2012

11. Elinkeinoelämän järjestöt: KKV:n ratkaisu hankintalain kiertämisestä tarpeellinen Suomen kasvulle: ”usko markkinatalouteen vahvistui” (May 2023), <https://ek.fi/ajankohtaista/tiedotteet/elinkeinoelaman-jarjestot-kkv-n-ratkaisu-hankintalain-kiertamisesta-tarpeellinen-suomen-kasvulle-usko-markkinatalouteen-vahvistui/> (in Finnish). Accessed 17 Oct 2023
12. Ghezzi, R.K., Korhonen, M., Vilpponen, H., Mikkonen, T.: The role of in-house procurement according to Finnish municipalities’ purchase invoice data. *Adv. Econ. Manage. Politic. Sci.* **17**(1), 373–378 (2023). <https://doi.org/10.54254/2754-1169/17/20231126>
13. Glaser, B.G.: *The Grounded Theory Perspective: Conceptualization Contrasted with Description*. Sociology Press, Mill Valley, CA (2001)
14. Government Procurement Manual 2017, In Finnish, Publications of the Ministry of Finance, vol. 29. Ministry of Finance (2017). <https://vm.fi/documents/10623/4040240/Valtion+hankintakasisikirja+2017/868b80fa-c2de-4328-ae93-36b17968f780/Valtion+hankintakasisikirja+2017.pdf?version=1.0>
15. Hartung, W., Kuźma, K.: In-house procurement-how it is implemented and applied in Poland. *Eur. Procure. Public Priv. Partnersh. Law Rev.* **13**(3), 171–183 (2018)
16. Hausmann, F.L., Quesner, G.: In-house contracts and inter-municipal cooperation-exceptions from the European union procurement law should be applied with caution. *Eur. Procure. Public Priv. Partnersh. Law Rev.* **8**, 231 (2013)
17. KKV vie Vantaan ja Keravan hyvinvointialueen Sarastia-hankinnan markkinaoikeuden arvioitavaksi, May 2023. <https://www.kkv.fi/uploads/sites/2/r-2022-60-1343.pdf> (in Finnish). Accessed 17 Oct 2023
18. Limited Liability Companies Act 624/2006 (2006). <https://www.finlex.fi/fi/laki/kaannokset/2006/en20060624>
19. Lith, P.: Tilastollinen raportti kunnallista yhtiöistä ja liikelaitoksista sekä niiden liiketoiminnan laajuudesta ja organisoinnista SoTe-uudistuksen kynnyksellä (May 2022), https://ek.fi/wp-content/uploads/2022/11/Hyvinvointialueliiketoiminta_2022_Pekka-Lith.pdf (in Finnish). Accessed 17 Oct 2023
20. Miemiec, W.: The application of in-house procurement by municipalities in municipal services management. In: *European Financial Law in Times of Crisis of the European Union*, p. 411 (2013)
21. Owners steer Tiera’s activities. Article in Finnish (2023). <https://tiera.fi/yritys/asiakasomistajat/>
22. Parking Brixen C-458/03. *Parking Brixen GmbH v Gemeinde Brixen and Stadtwerke Brixen AG*. ECJ, October 2005
23. *Sambre & Biesme SCRL and Commune de Farciennes v Société wallonne du logement (SWL)*. ECJ, December 2022
24. Sidosyksiköiden ulosmyynti. Selvityksen ensimmäisen vaiheen yhteenveto. In Finnish (June 2021), <https://www.kkv.fi/uploads/sites/2/2021/11/2021-kkv-selvityksia-6-2021-sidosyksikoiden-ulosmyynti-selvityksen-ensimmainen-vaihe.pdf> (in Finnish). Accessed 17 Oct 2023
25. Strauss, A., Corbin, J.: *Basics of Qualitative Research*. Sage Publications, Thousand Oaks (1990)
26. Urquhart, C.: Grounded theory for qualitative research: a practical guide. In: *Grounded Theory for Qualitative Research*, pp. 1–100 (2023)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Artificial Intelligence Procurement Assistant: Enhancing Bid Evaluation

Muhammad Waseem¹(✉), Teerath Das¹, Teemu Paloniemi¹, Miika Koivisto¹,
Eeli Räsänen¹, Manu Setälä², and Tommi Mikkonen¹

¹ Faculty of Information Technology, Jyväskylä University, Jyväskylä, Finland
{muhammad.m.waseem,dastzw,tommi.j.mikkonen}@jyu.fi,

{teemu.a.j.paloniemi,miika.j.koivisto,eeli.r.rasanen}@student.jyu.fi

² Solita, Tampere, Finland

manu.setala@solita.fi

Abstract. In modern business, maintaining competitiveness and efficiency necessitates the integration of state-of-the-art technology. This paper introduces the Artificial Intelligence Procurement Assistant (AIPA), an advanced system co-developed with Solita, a Finnish software company. AIPA leverages Large Language Models (LLMs) and sophisticated data analytics to enhance the assessment of procurement call bids and funding opportunities. The system incorporates LLM agents to enhance user interactions, from intelligent search execution to results evaluation. Rigorous usability testing and real-world evaluation, conducted in collaboration with our industry partner, validated AIPA's intuitive interface, personalized search functionalities, and effective results filtering. The platform significantly streamlines the identification of optimal calls by synergizing LLMs with resources from the European Commission TED and other portals. Feedback from the company guided essential refinements, particularly in the performance of ChatGPT agents for tasks like translation and keyword extraction. Further contributing to its scalability and adaptability, AIPA has been made open-source, inviting community contributions for its ongoing refinement and enhancement. Future developments will focus on extensive case studies, iterative improvements through user feedback, and expanding data sources to further elevate its utility in streamlining and optimizing procurement processes.

1 Introduction

Procurement bidding is a competitive process through which organizations seek to acquire goods, services, or projects from external suppliers or vendors [13]. This process involves inviting multiple suppliers to submit their proposals or bids for providing the required products or services [3]. The goal is to obtain the best value for the organization by allowing suppliers to compete based on factors such as cost, quality, delivery time, and other relevant criteria [10]. The

bidding process typically consists of the following steps: announcement or advertisement of the procurement opportunity, prequalification of potential suppliers, submission of bids, and the evaluation of bids [3].

Evaluating the bids in the bidding process needs reviewing the proposals to identify the supplier that aligns most closely with the organization’s requirements [5]. This evaluation considers various elements, including, quoted price, the quality of goods or services, the supplier’s prior history, their commitment, and other benefits they may offer. Evaluators depend on established benchmarks and rating mechanisms to fairly compare bids. The aim is to choose a proposal that not only fulfils the organizational needs but also provides the overall advantage and aligns with assessment standards [14].

The process of automated bidding evaluation leverages technology to enhance bid assessment within procurement procedures [7]. This technology-driven approach presents notable efficiency improvements, as automation significantly curtails the time and exertion required, thereby facilitating swift bid analysis [16]. Moreover, the automated systems introduce a crucial facet of uniformity in the application of evaluation criteria, effectively mitigating potential biases and errors that could arise [6]. This work stems from a research gap in the field – a need for streamlined, unbiased, and efficient bid assessment methods. In response to this research gap, we have developed and implemented the AIPA in collaboration with Solita Ltd¹. The development of AIPA marks a substantial stride in meeting the requisites for effective and impartial bid assessment. This system integrates LLMs with data analysis techniques, automating and elevating the entire bid evaluation process.

In the procurement, conventional manual bid assessment procedures often grapple with inadequacies. It is within this context that AIPA emerged, aiming to transcend the limitations of the status quo. Making adept use of LLMs, with ChatGPT taking center stage, AIPA swiftly comprehends intricate bid documents, applies predefined evaluation criteria, and distills crucial information for expedited human decision-making-whether to accept or reject proposals. One of AIPA’s distinctive strengths lies in its consistent application of evaluation criteria, eliminating subjective deviations. This stands in stark contrast to the inherent variability of manual evaluations, where individual interpretations can diverge significantly. Our industrial partners have expressed clear satisfaction with AIPA’s performance and capabilities.

In this paper, we are discussing the background in Sect. 2, followed by the proposed system in Sect. 3. The evaluation of the system is being presented in Sect. 4, and finally, we are concluding the study and suggesting future research in Sect. 5.

2 Background and Motivation

In modern business practices, procurement plays a key role in ensuring the acquisition of goods and services necessary for organizational operations [15]. Central

¹ <https://www.solita.fi/>.

to the procurement process is the critical task of bid evaluation, which involves assessing bids submitted by potential suppliers and selecting the most suitable ones based on a set of predetermined criteria [4]. However, traditional bid evaluation methods often face challenges related to subjectivity, manual effort, and potential bias, leading to inconsistencies and suboptimal decisions [11]. The use of Artificial Intelligence (AI) has brought about transformative changes in various industries, and procurement is no exception [9]. AI technologies have shown potential in automating and enhancing various aspects of the procurement process [8].

Machine learning enables software systems to learn from data patterns and make decisions based on specific requirements [2]. Models like GPT-3.5 and BERT have advanced the natural language processing, allowing machines to understand and create text that is similar to human [12]. These models have demonstrated their effectiveness in various tasks, including translating languages, generating text, answering questions, and analyzing sentiment [1].

Although AI in procurement is widely recognized, the area of bid assessment remains a critical area where AI based solutions could bring significant enhancements. Conventional bid evaluation methods often depend on manual analysis of bids, which can be time-consuming, labor-intensive, and subject to human biases [17]. Integrating LLMs into bid evaluation processes presents opportunities for organizations to enhance bid analysis, mitigate subjectivity, and improve the overall quality of decision-making.

Currently, bid evaluation methods are mostly characterized by manual efforts, extensive documentation, and the inherent risk of human-related errors. The need for more objective and efficient bid evaluation methods has become increasingly apparent, urging researchers and practitioners to explore novel avenues. In this context, our study aims to introduce an “Artificial Intelligence Procurement Assistant”. This tool uses the capabilities of LLMs, turning bid evaluation into a more efficient, objective, and informed process.

3 Proposed and Implemented System

AIPA is a system that we propose and implemented to streamline and enhance the procurement process for businesses. Leveraging AI capabilities, we have implemented a user-friendly and efficient way for users to find and assess relevant procurement notices from the European Commission’s TED portal. Our goal is to accelerate the procurement process by utilizing existing AI tools to assist businesses in making informed decisions about suitable procurement opportunities. Figure 1 present the key aspects of AIPA based on the high level system architecture diagram. Below, we provide a concise overview of AIPA’s key features.

- **User Interface (UI):** The AIPA UI serves as the primary point of interaction between users and the platform. Users, who are representatives of businesses, access the platform through this interface. We have implemented the UI to

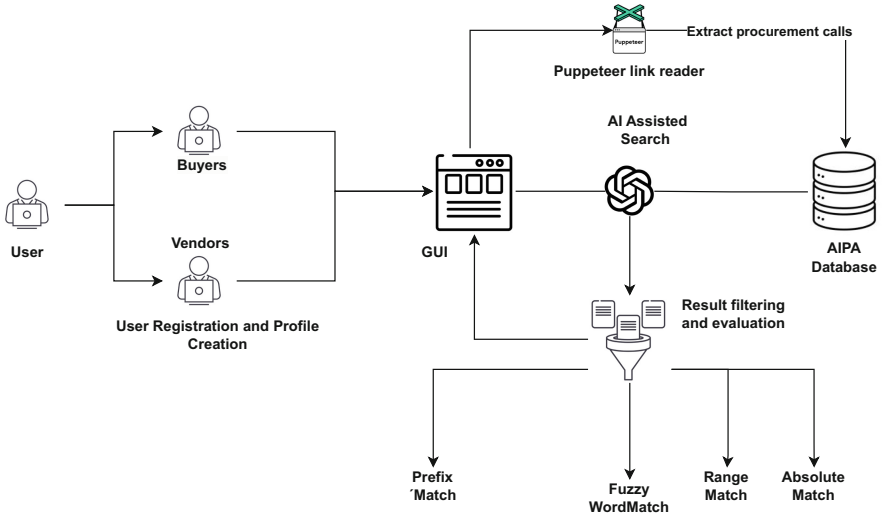


Fig. 1. High-Level System Architecture of AIPA

allow users to perform actions like registration, initiating searches, reviewing search results, and examining the generated list of procurement notices.

- **User Registration and Profile Creation:** A core functionality of AIPA is enabling users to register and create profiles. We have implemented an AI-assisted process to guide users in providing all the necessary parameters required for effective procurement notice searches. This AI-driven profile creation process enhances the relevance of search results and simplifies the registration process.
- **AI-Assisted Search:** AIPA utilizes ChatGPT to extract search parameters from user profiles. These parameters are then employed to conduct searches from AIPA database, which has been maintained to include procurement information from TED and other similar procurement websites. This is crucial for efficiently searching through large volumes of documents. Our implemented AI system, comprising multiple GPT agents with distinct roles and prompts to handles various tasks such as translation, keyword extraction, and generating similar words. These agents excel in distributed tasks rather than monolithic ones, contributing to improved results.
- **Results Filtering and Evaluation:** We have implemented a system that filters the search results obtained from the TED and other similar procurement websites and utilizes for evaluation. This process determines the relevance of these results to the user’s profile. By doing so, we ensure that the presented procurement notices align with the user’s business interests and preferences.
- **List Creation:** Based on the filtered and evaluated search results, our platform creates a list of the most suitable procurement notices. This list is presented to the user, providing a consolidated view of opportunities that match the user’s requirements.

- **ChatGPT Agents:** As a core of AIPA, we have integrated several ChatGPT agents for executing required tasks. These implemented agents assist in profile creation, parameter extraction, search execution, result evaluation, and justification generation. This component interacts with the TED portal to retrieve relevant procurement notices and performs AI-based analyses to enhance the overall quality of the procurement suggestions.

AIPA may act as a valuable resource for businesses seeking efficient and effective ways to navigate the complexities of procurement processes. By integrating ChatGPT seamlessly, we assist users in finding procurement opportunities that align with their specific needs, thereby simplifying and expediting the procurement journey.

4 AIPA Evaluation

The development of the AIPA system involved a partnership with Solita Ltd., critical for its testing and refinement. Solita Ltd. acted as the main evaluator and user, providing regular feedback during the development of AIPA.

Our teams worked together through weekly meetings and discussions, focusing on tailoring AIPA to meet user needs effectively. These interactions ensured that each feature developed was in line with what users expected and needed, with Solita Ltd. providing timely and essential feedback on every step.

Solita Ltd. was also key in assessing the main functions of AIPA. They tested how easy and effective the system was to use, including how users registered and searched within it. For example, they looked at how well the AI helped users set up their profiles and if this made search results more relevant.

They also examined AIPA's search feature, especially its ability to understand search terms and find the most appropriate results. The company checked the filtering options and made sure that the final list of procurement notices was what users were looking for.

Furthermore, they evaluated the ChatGPT agents incorporated into AIPA, particularly their role in translating languages, picking out key terms, and assessing search outcomes. Their real-world testing was essential for us to improve the system further.

To encourage others to contribute to AIPA's improvement, we made it open source on GitHub². This allows anyone interested to make changes and upgrades, helping AIPA to continue evolving and staying useful.

5 Conclusion

We have introduced the AIPA as an innovative solution aimed at streamlining and enhancing the procurement process for businesses in this paper. AIPA

² <https://github.com/koivupuu/AIPA>.

uses the power of AI, particularly ChatGPT, to provide a user-friendly and efficient platform for users to identify and evaluate relevant procurement opportunities. Through the development and implementation of AIPA, we have effectively addressed critical challenges encountered by businesses during traditional manual bid assessment procedures. AIPA has the potential to become an invaluable tool for businesses navigating complex procurement processes. By integrating ChatGPT, it simplifies and expedites procurement, assisting users in making informed decisions and improving overall efficiency. As AI continues to advance, AIPA's potential for enhancement and growth presents exciting opportunities for future research and development in the field of procurement assistance.

Looking ahead to further enhance AIPA, **future efforts** will first prioritize the refinement of its AI capabilities, conducting comprehensive case studies to evaluate real-world impacts, gathering user feedback to facilitate iterative improvements, broadening data sources, and exploring customization options. These endeavors will ultimately elevate its utility in streamlining and optimizing procurement processes.

References

1. Alberts, I.L., et al.: Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be? *Eur. J. Nucl. Med. Mol. Imaging* **50**(6), 1549–1552 (2023)
2. Bharadiya, J.P.: Leveraging machine learning for enhanced business intelligence. *Int. J. Comput. Sci. Technol.* **7**(1), 1–19 (2023)
3. Carayannis, E.G., Popescu, D.: Profiling a methodology for economic growth and convergence: learning from the EU e-procurement experience for central and eastern European countries. *Technovation* **25**(1), 1–14 (2005)
4. Cheaitou, A., Larbi, R., Al Housani, B.: Decision making framework for tender evaluation and contractor selection in public organizations with risk considerations. *Socioecon. Plann. Sci.* **68**, 100620 (2019)
5. Emilian, M.L.: Business-to-business online auctions: key issues for purchasing process improvement. *Supply Chain Manage. Int. J.* **5**(4), 176–186 (2000)
6. de la Fuente Garcia, S., Ritchie, C.W., Luz, S.: Artificial intelligence, speech, and language processing approaches to monitoring Alzheimer's disease: a systematic review. *J. Alzheimers Dis.* **78**(4), 1547–1574 (2020)
7. Geng, T., Sun, F., Wu, D., Zhou, W., Nair, H., Lin, Z.: Automated bidding and budget optimization for performance advertising campaigns. SSRN 3913039 (2021)
8. Guida, M., Caniato, F., Moretto, A., Ronchi, S.: The role of artificial intelligence in the procurement process: state of the art and research agenda. *J. Purchas. Supply Manage.* 100823 (2023)
9. Kulkov, I.: The role of artificial intelligence in business transformation: a case of pharmaceutical companies. *Technol. Soc.* **66**, 101629 (2021)
10. Li, S., Ragu-Nathan, B., Ragu-Nathan, T., Rao, S.S.: The impact of supply chain management practices on competitive advantage and organizational performance. *Omega* **34**(2), 107–124 (2006)
11. Liu, B., Yang, X., Huo, T., Shen, G.Q., Wang, X.: A linguistic group decision-making framework for bid evaluation in mega public projects considering carbon dioxide emissions reduction. *J. Clean. Prod.* **148**, 811–825 (2017)

12. Mohamadi, S., Mujtaba, G., Le, N., Doretto, G., Adjeroh, D.A.: ChatGPT in the age of generative AI and large language models: a concise survey. arXiv preprint [arXiv:2307.04251](https://arxiv.org/abs/2307.04251) (2023)
13. Roodhooft, F., Van den Abbeele, A.: Public procurement of consulting services: Evidence and comparison with private companies. *Int. J. Public Sect. Manage.* **19**(5), 490–512 (2006)
14. Scott, S.: Best-value procurement methods for highway construction projects, vol. 561. Transportation Research Board (2006)
15. Wang, G., Gunasekaran, A., Ngai, E.W., Papadopoulos, T.: Big data analytics in logistics and supply chain management: certain investigations for research and applications. *Int. J. Prod. Econ.* **176**, 98–110 (2016)
16. Watson, G.F., IV., Weaven, S., Perkins, H., Sardana, D., Palmatier, R.W.: International market entry strategies: relational, digital, and hybrid approaches. *J. Int. Mark.* **26**(1), 30–60 (2018)
17. Xueqing, W., Bingsheng, L., Allouche, E.N., Xiaoyan, L.: Practical bid evaluation method considering social costs in urban infrastructure projects. In: 2008 4th IEEE International Conference on Management of Innovation and Technology, pp. 617–622. IEEE (2008)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.


The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Platforms, Ecosystems and Data



Who Does What? Evolving Division of Responsibilities in a B2B Platform

Jaakko Vuolasto^(✉) 

LUT University, Lahti, Finland
jaakko.vuolasto@lut.fi

Abstract. To remain vital, a digital platform ecosystem requires governance. In the extant literature a platform ecosystem typically has a single focal actor who is responsible for the governance. We conducted a case study in heavy industry to understand how the responsibilities of a focal actor in governing a business-to-business platform ecosystem are shared and how they change. We observe the division of responsibilities and their changes as configurations. We conclude that the focal actor's responsibilities in a platform ecosystem are more multifaceted than the established view where a single actor has a stable set of responsibilities. The division of responsibilities in an ecosystem is subject to actor strategies and their positions in the supply chain. Thus, the strategic moves in an ecosystem are not made by a single actor but by multiple focal actors with multiple strategies.

Keywords: digital platforms · business-to-business · configurations · division of responsibilities

1 Introduction

Digital platforms are based on digital technologies and connectivity to utilize resources across company boundaries [1]. Different types of actors with varying degree of influence form a multi-sided market, a network where the actors are joined by contracts or other types of mutual dependencies [2]. A platform ecosystem is formed when the actors are organized around a platform [3]. This arrangement of actors requires governance: who has the power, who can make and what kind of decisions [4].

Most if not all this decision making is typically reserved for a single focal actor. This actor is referred as a platform owner [5], an orchestrator [6], or a keystone actor [7]. It has power over the ecosystem, especially the complementors that act in a certain niche within the ecosystem by extending the functionality of the platform [8, 9]. Ecosystems can also be decentralized in the sense that they have no single focal actor, such as in blockchain-based ecosystems [10]. However, we know little about the spectrum between these two extremes; how the governance responsibilities are given or taken in an ecosystem that is neither binary nor decentralized. This is especially relevant in business-to-business (B2B) platform ecosystems, where the relationships between the actors are different from the business-to-consumer context [11].

To fill this gap in research, we conducted a case study of a B2B platform ecosystem and its actors in a heavy industry with the following research question: *How are the responsibilities of a focal actor in a platform ecosystem shared?* To understand the division of responsibilities we interviewed different stakeholders and applied a configurational approach [12].

Our findings show that the division of responsibilities can be more multifaceted than the archetypical view presented in the platform literature. The focal actor's responsibilities are configurations and thus not stable but evolve over time, following actor relationships and interactions. The configurations reveal how the responsibilities of the focal actor in our case are divided between two actors. This increases our understanding of digital platform ecosystems especially in the B2B context that is more complex in terms of functionality [13] and stakeholders [14].

The rest of this paper is organized as follows. In Sect. 2 we present the responsibilities of a focal actor in a platform ecosystem and how they can be observed with a configurational approach. Section 3 describes our method. Our findings are in Sect. 4 and they are further discussed in Sect. 5. Finally, Sect. 6 concludes our work.

2 Background

2.1 Responsibilities of a Focal Actor in B2B Context

The responsibilities of an actor are linked with status and power. In a platform ecosystem the focal actor governs an ecosystem. Depending on the perspective this actor is recognized as a platform owner [1, 15, 16], leader [17], or an orchestrator [6, 18]. In our research we will use the term focal actor to refer to the central actor in the platform ecosystem.

The extant literature on the ecosystem actors and governance is vast. As our objective was to understand the responsibilities of a focal actor in a B2B context, we focused on the responsibilities that portray the characteristics of B2B platforms. Overall, the business models in B2B platforms are different compared to B2C [19]. They are manifested in different power relationships [11, 20] between the actors and in the responsibilities of the focal actor. The B2B context is considered more complex in terms of stakeholders [14] and supply chains [20]. The complexity is reflected in how the rules of a platform ecosystem are defined [21]. Typically the focal actor controls an ecosystem, by defining the rules in general [8, 15, 18] and also in respect to what the partners are allowed to do [1, 22]. However, the different business models of B2B can have an effect also on the defining of rules [19].

Platform creation requires laying the foundations for a nascent ecosystem [16]. It is the task of the focal actor to provide these foundations that the other actors build upon [6, 9]. This involves both technological decisions and architectural policies [23] suited for the B2B context, where the information systems are more complex [13].

Value co-creation and capture are in the heart of platform ecosystems, yet the mechanisms in the B2B context can be different from the B2C [16]. The focal actor not only seeks to extract value from the ecosystem, but it also shares value and resources [7]. This way, a focal actor is creating niches for the complementors [3, 7, 24]. The complementors add diversity and variability to the ecosystem by providing additional solutions

[1]. Their main incentive is the access to the customers of the platform provided by the focal actor [3]. This enables investments to a common future for the focal actor and its complementors [15, 17].

As the largest group of actors the end-users are the source of the financial value in platform ecosystems [3, 6]. In addition to creating niches, the focal actor is in charge of attracting end-users and facilitating interactions between the complementors and the end-users [15, 25]. It is the focal actor that provides the complementors with access to the customer base of the platform ecosystem [3, 7, 24]. The key responsibilities of a focal actor are summarized in Table 1 below.

Table 1. Summary of the focal actor's responsibilities.

Key Responsibilities	Literature
Defining the rules: who can participate and what the participants are allowed to do	Cenamor, 2021; Manikas & Hansen, 2013; Tiwana, 2013; Gawer, 2020; Ruippo et al., 2023; Ritala & Jovanovic, 2024
Laying foundations: technological and / or architectural principles	Ghazawneh & Henfridsson, 2013; Jansen, 2020; Hodapp et al., 2019; Karhu et al., 2020; Foerderer, 2019
Niche creation for complementors by sharing resources and value	Iansiti & Levien, 2004; Jacobides et al., 2018; Williamson & De Meyer, 2012; Hodapp et al., 2019; Moore, 1993; Cenamor, 2021
Attraction: both end-users and complementors	Cenamor, 2021; Eisenmann et al., 2009; Pauli et al., 2021
Access granting: complementors to the customer base	Jacobides et al., 2018; Moore, 1993; Iansiti & Levien, 2004; Williamson & De Meyer, 2012

2.2 Configurational Approach to Responsibilities

In the existing research the focal actor is depicted as a single entity that is exclusively responsible for its own key tasks; respectively, the complementors are solely responsible for their tasks [1, 3, 8]. These responsibilities are presented rather stable, there is very little or no room for variance or dynamics. However, the complexity and specifics of the B2B context [11, 19] call for a broader perspective. Viewing the focal actor's responsibilities as a configuration can extend our understanding of B2B platform ecosystems. A configuration consists of characteristics or elements that occur together and align into patterns [12, 26]. The elements of a configuration are interdependent and an orchestrating theme connects them [27]. Importantly, a configuration is dynamic, it can change over time [27].

Configurations have been applied in analyzing the adoption of inter-organizational information systems [28], where the configuration consists of five elements: organizing vision, key functionality, structure, mode of interaction, and mode of appropriation. There are configurational studies also in platform research, for instance [29]. However,

it has not been used extensively although the features of configurational approach such as emergence and equifinality [26] make it suitable for this purpose.

Configurations emerge from the strategies the actor implements [26]. In the platform context Eisenmann et al. [25] portray two types of strategies for a focal actor. A horizontal strategy allows other actors to participate in the commercialization and technical development of the platform, even broadening the sponsorship to other actors by giving them access to the development of the core technology. A vertical strategy on the other hand contains decisions for example on the extent of complementor access to the platform and make-or-buy decisions: whether the focal actor should include functionality provided by complementors into the platform core. Another way to view the strategies of a focal actor is with a keystone or a dominator perspective [7]. In a keystone strategy an actor focuses on the external resources and occupies only a limited number of nodes in an ecosystem. A dominator strategy is opposite in the sense that it aims at both value creation and capture, thwarting the creation of alternative solutions by other companies. We focus on the configuration of the responsibilities of a focal actor in the B2B context, and the strategies they are based on.

3 Research Method

We conducted a case study to investigate the responsibilities of a focal actor in a B2B platform ecosystem. Aiming to understand a contemporary phenomenon in its real-life environment with a “how” question justified our selection of the research method [30]. A case study should offer something new and a basis for analytic generalization by shedding “empirical light on some theoretical concepts or principles” [30]. We selected wood supply in Finland as our case because it presented a combination of maturity and novelty. A digital platform connects groups of heterogenous actors and their information systems, forming an ecosystem. There are competing wood buyer companies that purchase timber from the forest owners and outsource the harvesting operations to smaller contractor companies. In their operations the contractors utilize forest machines provided by machine manufacturers. Both the wood buyers and the contractors rely heavily on information systems provided by different vendors. The introduction of the platform transformed the information systems landscape. This setting provides a novel view to focal actors in a B2B context: not a single incumbent company but neither a completely decentralized ecosystem. Using the configurational approach that explores holistically the “why” and “how” aspects guided us in understanding the context [27].

The information systems in wood supply were in two categories: the enterprise resource planning (ERP) systems of the wood buyers and the control systems in the forest machines. The control systems depend on the data provided by the ERP systems, and they send the data about performed work back to the ERP systems. Previously the two types of systems had been connected directly to each other. In 2013 three large wood buyer companies (Founders from here on) started a joint effort. Instead of company-specific development they chose to implement a digital platform that would cover a share of functionality that had been in the ERP systems. This forestry platform (FPF) and its functionality were aimed mostly at the contractors. The Founders selected a software company (SoftwareCo from here on) and outsourced the implementation and

operation of the FPF to it. FPF went operational in 2016 and by 2019 the Founders had all their operations on the platform.

Our case study protocol was designed in early 2021, including the data sources, informed consent, interview questions, and a timeline for the research [30]. In the beginning, the extant literature gave us the first frame of reference for a focal actor's responsibilities [2, 3, 9]. Our primary data source consisted of 31 interviews conducted by the first author in 2021. The interviewees were selected to cover the variety of actors in the FPF ecosystem: decision makers and subject matter experts working in wood buyer companies, different types of contractor companies, machine manufacturers, and representatives of SoftwareCo. In reaching out to the interviewees we relied partially on the first author's prior working experience in SoftwareCo, which helped establish contacts and provided a common language. The interviewees, their organizations and roles are described in Table 2.

Table 2. List of interviewed companies and persons.

Organization	Interviewees and their roles
Consultancy services for the founders	Project Manager (9)
Contractor 1	Account Manager (2)
Contractor 2	Manager (11)
Contractor 3	Manager (12)
Contractor 4	CEO (13)
Contractor 5	CEO (14)
Contractor 6	CEO (16)
Contractor 7	CEO (19)
Educational Institution	Teacher, Harvesting (20)
Global Machine Manufacturer A	Technical Customer Support Manager (22)
Global Machine Manufacturer B	Product Group Manager (31)
SoftwareCo operating the platform and providing enterprise systems	Product Owner (17); Service Manager (21); Service Manager (23); Product Owner (25); General Manager (26); Key Account Manager (27)
State-funded organization for forestry	Specialist (30)
Wood Buyer A: Founder	Senior Vice President, Development (3); ICT Solution Designer (10)
Wood Buyer B: Joined later	System Specialist (4)
Wood Buyer C: Founder	Development Manager (5); Development Specialist (6); Team Lead, Information Management (8)

(continued)

Table 2. (continued)

Organization	Interviewees and their roles
Wood Buyer D: Founder	SVP, Innovation and Development (7); Solution Architect (15); Development Manager, Harvesting (28); Operations Manager (29)
Wood Buyer E: Joined later	Manager (18)
Wood Buyer F: Joined later	Manager (24)
Wood supply R&D company	CEO (1)

The interview questions were grouped into four themes: the beginning and the idea behind FPF, day-to-day operation, development, and the community around FPF. The interview questions are available at <https://bit.ly/40q6Q5X>. The first author conducted the interviews remotely. The interviews were recorded and transcribed, and the Atlas.TI software was used in the analysis of the transcripts. We analyzed the interview data by the principles of grounded theory [31]. We started with initial codes that identified the responsibilities of each actor in the ecosystem as perceived by the interviewees. During the analysis the position and responsibilities of a focal actor were quite often attributed to the Founders and the SoftwareCo. Thus, we strived to get a comprehensive data set from these actors.

When no new responsibilities emerged from the data, we had reached conceptual saturation and continued the analysis by looking at the context and process [31]. There was a pattern in how the responsibilities of each actor were perceived – by an actor itself but also by others. This pattern deviated from the established view in platform literature. Also, the emerging pattern clearly changed over time: first the Founders were perceived to be the focal actor, but later the responsibilities of the focal actor became shared. We then returned to seminal works on the responsibilities of the focal actor to compare our findings with the literature. The concept of configuration [12] helped us in understanding the patterns in the division of responsibilities and their development, rooted in different types of strategies.

4 Findings

4.1 Actors in Forestry Platform

The FPF ecosystem has five groups of actors: the wood buyers, the companies that provide ERP systems for the wood buyers, contractors, machine manufacturers, and SoftwareCo that implements and operates FPF. The actors are shown in Fig. 1. SoftwareCo has formal agreements on the use of FPF with the contractors and wood buyers. Machine manufacturers provide the forest machines and the control systems to the contractors, and respectively the ERP providers provide the enterprise systems for the wood buyers. SoftwareCo competes to some extent with both the machine manufacturers and ERP providers. Although no formal agreements exist between the machine manufacturers and wood buyers, the relationship is important to both actors.

In its core FPF contains applications for forestry operations and interfaces for the wood buyer ERP systems and the control systems in the forest machines. When a wood buyer purchases wood from a forest owner, the ERP system of the wood buyer provides the data to a specific contractor, via FPF core. The contractor then plans the harvesting operations: when and by which machine. This planning takes place in the application belonging into FPF core. Once the planning is completed, the data for the working sites is transferred to the forest machine and into the control system. During and after the harvesting operations the control system provides data about the amount and quality of the wood harvested. This data travels via FPF core back to the ERP system of the wood buyer.

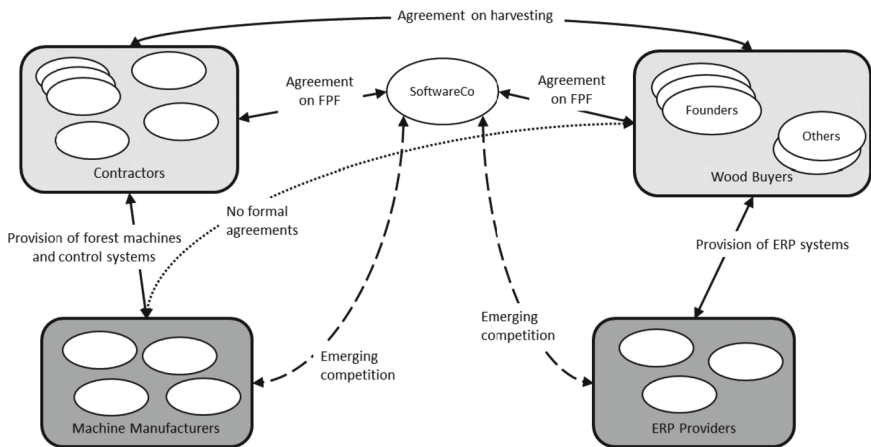


Fig. 1. Actors in FPF ecosystem.

The wood buyers' main objective is to secure a stable flow of the raw material. They purchase wood from the forest owners and outsource the harvesting operations to their contractors. A contractor has an agreement with one or more wood buyers, and the wood buyers have substantial negotiating power over their contractors. Using FPF is obligatory for the contractors. SoftwareCo is an actor with considerable amount of power and a strong presence in the ecosystem. In addition to running and developing FPF core SoftwareCo also provides ERP systems for one of the Founders and other wood buyers that joined FPF later.

4.2 From Common Problem Scope to Assembly Configuration

In what follows we show the development of the division of responsibilities through two different configurations. First, the *Assembly* configuration refers to the design and creation of FPF, where the Founders have the key responsibilities. It is followed by the *Established* configuration, where the responsibilities are shared. The overall change is described in Fig. 2.

The Founders shared a need for major renewal of their enterprise systems. This problem was not merely about a major upgrade to information systems but about developing

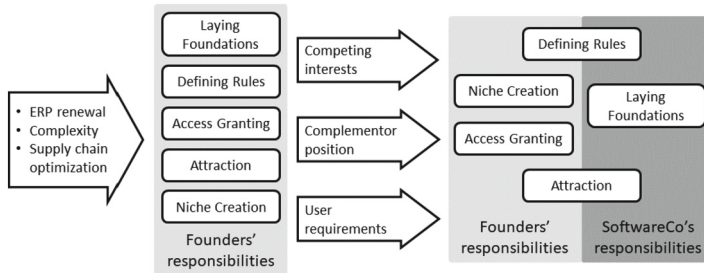


Fig. 2. The overall development of the focal actors' responsibilities.

new solutions to common problems. Although competing, they found a common area of interest in collective supply chain optimization: “*we have to find a common tool across firm boundaries for steering and planning the [contractor] work for multiple wood buyers*” (interviewee #7). The effects of having to use multiple, company-specific information systems had affected the contractors the most: “*each [wood buyer] company had their dedicated systems and if a contractor worked for more than one wood buyer, then there were multiple parallel systems in a single forest machine*” (interviewee #3). Also, the machine manufacturers suffered from the complexity of the situation: “*whenever we delivered a new or used machine, there was a maximum of 14 different [wood buyer] systems to install*” (interviewee #22).

The Founders identified the common functionality and designed it to be the core of a new platform. In 2012 they engaged in a shared sponsorship of a future platform and decided to outsource the implementation. The outsourcing to SoftwareCo acted as a value co-creation and sharing activity. The Founders designed the business model so that the revenue was to be collected by SoftwareCo: “*the agreements were made so that [SoftwareCo] owns the software and part of the business model is that the company gets compensated for providing the service*” (interviewee #3). An exclusive access to the customer base was granted for SoftwareCo. With these actions the Founders aligned interests with SoftwareCo.

The Founders defined a framework for both the architecture and the governance of the platform ecosystem. The former was materialized in the design specifications of the platform, including the principles for how the complementing solutions could and should extend the platform core. The latter, a governance framework, included rules for other organizations to join the platform, rules for the common development, and rules for the future service provider in the form of a service level agreement. There was no need to attract end-users since the wood buyers made it mandatory for their contractors to use the platform.

The Founders did not at this point create a technological core to extend, but they designed the first niche by outsourcing the technical specification and implementation to SoftwareCo. With respect to the machine manufacturers, the Founders designed a niche for them as well but left the scope vaguer. The aim was at a semi-open ecosystem, based on an international standard, but no criteria for value sharing with the manufacturers were defined. Yet due to the position of the Founders and the strategy of the manufacturers,

the interests were aligned enough, and the machine manufacturers adapted to the major market change initiated by the Founders.

The development of FPF started in 2013 and led to the first deployments in 2016. We have identified the division of responsibilities in this phase as the Assembly configuration of the platform. The Assembly configuration reflected the strong position of the Founders; they had all the key responsibilities as displayed in Table 3. They financed the design and implementation of FPF, being the only source of financial value in the ecosystem. The ERP providers and machine manufacturers were complementors. At this point SoftwareCo was positioned as a complementor instead of a focal actor. It started from a niche created by the Founders, and it had to operate by the rules defined by the Founders. Also, the Founders had the power to grant SoftwareCo the access to all of their contractors.

4.3 Reaching the Established Configuration

By 2019 all the Founders were using the platform. As the platform gradually reached an established position in terms of installed base and the stability of operations, the initial problems were solved. The platform was a tool that served the actors in a fashion that was perceived good enough. From the wood buyer point of view, it was considered irreversible: *“the way I see it [FPF] is here to stay”* (interviewee #1).

Because the use of the platform was mandatory for contractors, whenever a new contractor started to work for a wood buyer, it also became a customer of SoftwareCo. However, these additions were relatively small, which made SoftwareCo to search for growth by bringing new wood buyers to the platform ecosystem. To reach the goal SoftwareCo bundled FPF and its deployment with enterprise systems it provided: *“[FPF] is a part of our service offering for managing the entire value chain in wood supply, ... in a sense one module of the overall solution”* (interviewee #26).

In this way SoftwareCo gradually moved toward being a focal actor but at the same time held on to the complementor niche as an ERP provider. As a result of this bundling, between 2019 and 2021 several new wood buyers started the use of FPF. The installed base of the platform grew in bursts. However, this bundling based on a dominator strategy meant that the development resources of SoftwareCo were allocated in a different way compared to the previous configuration. The Founders perceived that they did not get as much development resources as was agreed. Although the interests of the two actors had been aligned, they now started to deviate.

With the platform core implemented, SoftwareCo was responsible for providing the technological and architectural foundations. The company also took part in defining the rules, especially regarding what the other actors were allowed to do. It had identified the machine manufacturers as a source of possible competition and wanted to keep them at an arms-length distance. The control system and its interaction with FPF constituted an example of how external systems extend the functionality provided by the platform core. However, the manufacturers' software offering contained also features that were competing with some of the functionality present in the platform core.

The contractors acknowledged that the platform was implemented, but not complete. In addition to interoperability with machine manufacturers' solutions, another area where significant needs for improvement prevailed was in the planning of contractor operations.

The issues were rooted in the autonomy given to the contractors. It had led to a situation where the operating volumes of contractor companies had grown, sometimes causing performance issues in the platform core, as described by interviewee #11: “*now that the amount of working sites has reached thousands, the system is lagging, quite regularly*”. These issues were reported both to the wood buyers and SoftwareCo but solving them was progressing slowly.

At this point there were multiple problems: the machine manufacturers’ position as complementors, addressing the emerging needs of the contractors, and serving the Founders as well as new wood buyers. The platform was no longer only an initiative of the Founders but nor was it completely governed by SoftwareCo. It was not easy to achieve an alignment among the Founders, SoftwareCo, and the other actors other. The Founders held on to the principles inscribed in the governance framework of the platform. SoftwareCo argued that it had fulfilled the obligations and as a focal actor took steps in defining the rules and attracting new users. The tensions led gradually to a new division of responsibilities, which we identified as the *Established configuration*, presented in Table 3. The bolded responsibilities indicate a change compared to the Assembly configuration.

Table 3. The division of responsibilities in the two configurations.

Key Responsibility	Responsible Actor in the Assembly Configuration	Responsible Actor in the Established Configuration
Defining rules	Founders	Founders and SoftwareCo
Laying foundations	Founders	SoftwareCo
Attraction	Founders	Founders and SoftwareCo
Niche creation	Founders	Founders
Access granting	Founders	Founders

A clear shift was in how the provision of technological and architectural foundations was now completely SoftwareCo’s responsibility. Modifications to the platform core and to the interfaces were designed and implemented by the company. All actors recognized and accepted this.

Setting the rules was divided between the Founders and SoftwareCo. Aligning the interests in respect to machine manufacturers’ position serves as an example. The manufacturers had recognized the need to strengthen their position in the ecosystem. They were interested in enriching their solutions with the data in the platform core and even using their applications instead or side by side with the core applications provided by SoftwareCo. However, SoftwareCo was reluctant to give them a bigger role and acted cautiously, avoiding any moves that would weaken its position. Instead, SoftwareCo focused on serving the Founders and attracting new wood buyers.

The discussion about exchanging data between FPF core and control systems had been going on since 2020, but with little progress. Manufacturers recognized SoftwareCo as a focal actor, but they also understood the Founders’ fundamental role: “*it is a wood*

buyer solution for transferring data to and from the forest machines. I see it primarily as a wood buyer effort" (interviewee #31). Some of the larger manufacturers asked the Founders to help in the negotiations with SoftwareCo. The Founders used their power in aligning the interests of the manufacturers, SoftwareCo, and the contractors. The argument that FPF was developed primarily for the contractors was interpreted so that the obligatory use of the platform should not block the use of additional applications provided by machine manufacturers: *"if a contractor wants to buy a fit solution from a machine manufacturer, it should be allowed and [FPF] should not block it"* (interviewee #28). Furthermore, SoftwareCo was not in the position to grant or deny the manufacturers the access to the customer base, because manufacturers already had contractors as their customers.

In summary, the Founders initiated the FPF development. First, in the Assembly configuration the Founders had all the responsibilities of the focal actor. Additionally, the Founders acted also as end-users. SoftwareCo was positioned as a complementor in the ecosystem. Later, in the Established configuration the focal actor's responsibilities were shared across the Founders and SoftwareCo. The Founders' position in the supply chain gave them power over their contractors, and as the creators of FPF their views carried more weight over the other wood buyers that joined later. However, there was no single focal actor that governed the ecosystem at all times.

5 Discussion

5.1 Shared Responsibilities and Multiple Strategies

We studied the focal actors and their relationships in a platform ecosystem to understand the division of responsibilities. In the literature a focal actor is considered to have power over the ecosystem and complementors due to one-to-many structure and asymmetric dependencies [32]. We provide a new perspective in understanding the early phases of a platform development [33]. Our research shows that there is an overall division of responsibilities in an ecosystem, a configuration of actors and their responsibilities that changes over time [12]. The configurational approach has been used in information systems adoption [28] but only scarcely in the platform research [29].

Although configurations open up a space of possibilities, not all configurations are likely or even possible [27]. The view that focuses on a single focal actor with fixed responsibilities is the prevailing in the extant literature [4, 17, 18]. Our findings indicate that another configuration is possible. In a classic platform ecosystem a focal actor would solve governance issues [9, 23]. In other words, a focal actor would play the main role [34]. When the ecosystem is complex, a single focal actor can be absent [14] or an ecosystem can also be completely decentralized [10]. The FPF ecosystem presents another option where there is no single focal actor nor is the ecosystem completely decentralized. The focal actor's responsibilities in FPF ecosystem are divided between two actors, which can be viewed as an example of power dynamics in the B2B context [11]. B2B platforms include matchmaking, marketplaces, and supply chains as well [19, 20]. If a B2B company wants to succeed with a digital platform it should acknowledge that there are lessons to be learned from successful B2C companies. At the same time it is important to recognize that not all the B2C strategies are applicable to B2B network [35].

Our findings show how a platform creation can be a joint effort. In this effort, defining the responsibilities of different actors is a crucial task. Ensuring sufficient alignment of interests is a critical success factor [34].

The Founders implemented a horizontal platform strategy by allowing other wood buyers to join the platform [25]. Their approach was close to keystone strategy where an actor does not dictate an ecosystem [7]. However, this openness was directed toward other wood buyers. With respect to the contractors, the Founders did dominate. This was due to the contractual relationship and the supply chain. Joining the platform is easy for a contractor but leaving is not an option as long as it works for a wood buyer using FPF. This helped in aligning the interests of the two focal actors [34].

When the focal actor responsibilities became shared in the Established configuration, SoftwareCo started to utilize dominator strategy, aiming to occupy several niches in the ecosystem [7]. SoftwareCo bundled its offerings, providing a solution for the complete value chain [33]. Because the market was limited, SoftwareCo utilized a vertical platform or even a product strategy to search for growth [25]. The vertical strategy was utilized also with respect to the machine manufacturers. The emerging competition called for balancing the different strategies and tactics [23]. As the focal actor role was shared, there was no single owner or a focal actor that could decide the level of openness [25]. The Founders had to take a role in seeking the balance, for the overall health of the ecosystem [7]. The arrangement of two focal actors was relatively stable. However, the diversity of the complementing solutions in the ecosystem remained limited, due to limited number of complementors [19] and the tension between SoftwareCo and the machine manufacturers. Whereas the tension between a focal actor and a complementor is characterized in the literature as delicate [9], in FPF ecosystem it was overpowering, causing stagnation in the relationship between SoftwareCo and machine manufacturers.

While the literature presents a framework for decision making where focal actor decides platform strategies and complementor niches [5, 33], it can be so that the choice of strategies is not for a single actor to make. Some decisions may also require a regulator [1]. The extant literature does not include a regulator in the actors of a platform ecosystem, although the impact of regulation can be significant [36].

5.2 Limitations and Future Research

As our work was qualitative research, concerns for validity cannot be removed absolutely [31]. We briefly describe the actions taken to mitigate descriptive, interpretive, and theoretical validity. Our interviews were recorded and transcribed to improve descriptive validity. The first author was also responsible for the coding and analysis. This way the overall content of an interview, including contextual information recognized by the researcher was available. For interpretive validity, identifying the participants' perspective of events is crucial. To foster this goal, the data collection was extensive, aiming at data triangulation [30]. The first author's familiarity with the domain provided common language and mutual understanding in the interviews. Regarding theoretical validity, the configurations we identified are not likely the final ones. The configurational approach allows for the variation in order and reassessments of configurations [26, 27], thus leaving room for seeking alternative explanations [30]. By using configurational approach, we strived for utilizing a theory that would validate our research. This provides starting

points for future research in the B2B context, including the actors' responsibilities more generally, and the role of a regulator in a platform ecosystem.

6 Conclusion

We presented an alternative approach to view the division of responsibilities in a platform ecosystem, based on a case study of a B2B digital platform in a heavy industry and utilizing configurations as the theoretical framework. From the extant literature we collected responsibilities especially relevant in a B2B context that defined the archetypical division of responsibilities. Our findings suggest that the allocation of responsibilities is more multifaceted than the archetypical setting where a single focal actor has a stable set of responsibilities. There is variety in how the responsibilities are allocated – the actors' responsibilities are configurations and thus not stable but evolve over time, following actor relationships and their strategies. The configurations revealed the focal actor's role that was divided between two actors. As there was no single actor that steered the platform ecosystem, there was no single strategy but a combination of many. The shared role of a focal actor was a potential source of confusion but also a factor that stabilized the platform ecosystem.

References

1. Gawer, A.: Digital platforms' boundaries: the interplay of firm scope, platform sides, and digital interfaces. *Long Range Plann.* **54**, 102045 (2020). <https://doi.org/10.1016/j.lrp.2020.102045>
2. Tan, B., Pan, S., Lu, X., Huang, L.: The role of IS capabilities in the development of multi-sided platforms: the digital ecosystem strategy of Alibaba.com. *J. Assoc. Inf. Syst.* **16**(4), 248–280 (2015). <https://doi.org/10.17705/1jais.00393>
3. Jacobides, M.G., Cennamo, C., Gawer, A.: Towards a theory of ecosystems. *Strateg. Manag. J.* **39**, 2255–2276 (2018). <https://doi.org/10.1002/smj.2904>
4. Tiwana, A., Konsynski, B., Bush, A.A.: Research commentary—platform evolution: coevolution of platform architecture, governance, and environmental dynamics. *Inf. Syst. Res.* **21**, 675–687 (2010). <https://doi.org/10.1287/isre.1100.0323>
5. Cenamor, J., Frishammar, J.: Openness in platform ecosystems: innovation strategies for complementary products. *Res. Policy* **50**, 104148 (2021). <https://doi.org/10.1016/j.respol.2020.104148>
6. Jansen, S.: A focus area maturity model for software ecosystem governance. *Inf. Softw. Technol.* **118**, 106219 (2020). <https://doi.org/10.1016/j.infsof.2019.106219>
7. Iansiti, M., Levien, R.: *The Keystone Advantage: What the New Dynamics of Business Ecosystems Mean for Strategy, Innovation, and Sustainability*. Harvard Business Review Press, Boston (2004)
8. Tiwana, A.: *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Newnes (2013)
9. Ghazawneh, A., Henfridsson, O.: Balancing platform control and external contribution in third-party development: the boundary resources model. *Inf. Syst. J.* **23**, 173–192 (2013). <https://doi.org/10.1111/j.1365-2575.2012.00406.x>
10. van Pelt, R., Jansen, S., Baars, D., Overbeek, S.: Defining blockchain governance: a framework for analysis and comparison. *Inf. Syst. Manage.* **38**, 21–41 (2021). <https://doi.org/10.1080/10580530.2020.1720046>

11. Hasler, D., Krummy, B., Schallmo, D.: Characteristics of digital platforms from a b2b perspective – a systematic literature review. In: PACIS 2022 Proceedings (2022)
12. Miller, D.: Configurations of strategy and structure: towards a synthesis: summary. *Strateg. Manage. J.* **7**(233), 1986–1998 (1986)
13. Foerderer, J., Kude, T., Schuetz, S.W., Heinzl, A.: Knowledge boundaries in enterprise software platform development: antecedents and consequences for platform governance. *Inf. Syst. J.* **29**, 119–144 (2019). <https://doi.org/10.1111/isj.12186>
14. Petrik, D., Herzwurm, G.: Towards the iiot ecosystem development - understanding the stakeholder perspective. In: ECIS 2020 Res. Pap. (2020)
15. Cenamor, J.: Complementor competitive advantage: a framework for strategic decisions. *J. Bus. Res.* **122**, 335–343 (2021). <https://doi.org/10.1016/j.jbusres.2020.09.016>
16. Hodapp, D., Hawlitschek, F., Kramer, D.: Value co-creation in nascent platform ecosystems: a Delphi study in the context of the internet of things. In: ICIS 2019 Proceedings (2019)
17. Moore, J.F.: *Predators and Prey: A New Ecology of Competition* (1993)
18. Manikas, K., Hansen, K.M.: Software ecosystems—a systematic literature review. *J. Syst. Softw.* **86**, 1294–1306 (2013)
19. Ritala, P., Jovanovic, M.: Platformizers, orchestrators, and guardians: three types of B2B platform business models. In: Aagaard, A. and Nielsen, C. (eds.) *Business Model Innovation: Game Changers and Contemporary Issues*. Palgrave Macmillan (2024)
20. Anderson, E.G., Lopez, J., Parker, G.G.: Leveraging value creation to drive the growth of B2B platforms. *Prod. Oper. Manag.* **31**, 4501–4514 (2022). <https://doi.org/10.1111/poms.13866>
21. Jovanovic, M., Sjödin, D., Parida, V.: Co-evolution of platform architecture, platform services, and platform governance: expanding the platform value of industrial digital platforms. *Technovation* **118**, 102218 (2022). <https://doi.org/10.1016/j.technovation.2020.102218>
22. Ruippo, A., Koskinen, K., Rossi, M.: Actor Positioning and Its Implications to Value Co-Creation in SaaS Ecosystems (2023)
23. Karhu, K., Gustafsson, R., Eaton, B., Henfridsson, O., Sørensen, C.: Four tactics for implementing a balanced digital platform strategy. *MIS Q. Exec.* **19**, 105–120 (2020). <https://doi.org/10.17705/2msqe.00027>
24. Williamson, P.J., De Meyer, A.: Ecosystem advantage: how to successfully harness the power of partners. *Calif. Manage. Rev.* **55**, 24–46 (2012). <https://doi.org/10.1525/cm.2012.55.1.24>
25. Eisenmann, T., Parker, G., Van Alstyne, M.: Opening platforms: how, when and why? In: *Platforms, Markets and Innovation*. pp. 131–162. Edward Elgar Publishing (2009)
26. Meyer, A.D., Tsui, A.S., Hinings, C.R.: Configurational Approaches to Organizational Analysis. *Acad. Manage. J.* **36**, 1175–1195 (1993). <https://doi.org/10.2307/256809>
27. Miller, D.: Configurations Revisited. *Strateg. Manag. J.* **17**(505), 1986–1998 (1996)
28. Lyytinen, K., Damsgaard, J.: Inter-organizational information systems adoption – a configuration analysis approach. *Eur. J. Inf. Syst.* **20**, 496–509 (2011). <https://doi.org/10.1057/ejis.2010.71>
29. Fuerstenau, D., Auschra, C., Klein, S.: A configuration approach to multi-sided platforms in healthcare: an ALS platform case. In: 41st International Conference on Information Systems (ICIS) 2020, Hyderabad, India, 13 December 2020
30. Yin, R.K.: *Case Study Research and Applications: Design and Methods*. SAGE, Los Angeles (2018)
31. Corbin, J., Strauss, A.L.: *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE, Los Angeles (2015)
32. Hurni, T., Huber, T.L., Dibbern, J.: Power dynamics in software platform ecosystems. *Inf. Syst. J.* **32**, 310–343 (2022). <https://doi.org/10.1111/isj.12356>
33. Schermuly, L., Schrieck, M., Wiese, M., Krcmar, H.: Developing an industrial IoT platform – trade-off between horizontal and vertical approaches. In: *Wirtsch 2019 Proceedings* (2019)

34. Sarker, S., Sarker, S., Sidorova, A.: Understanding business process change failure: an actor-network perspective. *J. Manag. Inf. Syst.* **23**, 51–86 (2006). <https://doi.org/10.2753/MIS0742-1222230102>
35. Riemensperger, F., Falk, S.: Three Lessons From Germany's Platform Economy, <https://sloanreview.mit.edu/article/three-lessons-from-germanys-platform-economy/>. Accessed 25 Nov 2022
36. Bazarhanova, A., Yli-Huumo, J., Smolander, K.: Love and hate relationships in a platform ecosystem: a case of Finnish electronic identity management. In: Proceedings of the 51st Hawaii International Conference on System Sciences, pp. 1493–1502 (2018). <https://doi.org/10.24251/HICSS.2018.187>




Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Understanding User Feedback in Software Ecosystems: A Study on Challenges and Mitigation Strategies

Bachan Ghimire^(✉) , Ze Shi Li , and Daniela Damian 

University of Victoria, British Columbia, Canada
{bachan48,lize,danielad}@uvic.ca
<http://thesealgroupp.org>

Abstract. Online user feedback has become an essential mechanism for software organizations to gain insight into user concerns and to recognize areas for improvement. In software platform ecosystems, staying abreast of user feedback is particularly challenging due to the multitude of feedback channels and the complex interplay with third party applications. In this paper we report from a mixed-method study of user feedback from over 40,000 relevant reviews from 139 SECO platforms out of 2.4 million online user reviews scraped from 283 retrieved SECO platforms. Through thematic analysis and machine learning classifiers with high accuracy, we identified and analyzed six categories of user challenges in the areas of Integration, Customer Support, Design & Complexity, Privacy & Security, Cost & Pricing, and Performance & Compatibility. Our analysis also shows a significant growth of SECO user feedback in the past five years, highlighting the importance of understanding such user feedback as well as research methodologies to automatically study online user concerns in software ecosystems. To further understand mitigation strategies for challenges reported by end users, we interviewed four executives from large ecosystems and describe strategies in addressing those identified challenges. This research is a first large scale study of user feedback in software ecosystems; the categories of user concerns are hopefully useful in guiding platforms in designing and fostering better software ecosystems. Our methodology for automatically classifying the user feedback that is SECO-related can also serve as guidance for future studies that can further advance our understanding of user feedback and how to integrate it into improved software ecosystems.

Keywords: software ecosystem · machine learning · user feedback

1 Introduction and Background

Over the last decade, there has been a significant change in the way software companies function and use platforms as a type of open innovation to expand their markets and stakeholders, and have seen a significant increase in software

usage. These platforms serve as the foundation for creating software ecosystems (SECO)s, where the platform provider, also known as the keystone organization, collaborates and innovates with other software vendors [1, 2]. Software ecosystems are complex and dynamic systems, consisting of various software components, platforms, and developers that interact with each other [1]. Companies such as HubSpot, Salesforce, Xero, Slack, Shopify, and Wix have thrived from their integration, marketplace, innovation, and other qualities that make a thriving ecosystem [3].

Various operating system-specific application stores, marketplaces, public review websites, and keystone platforms like Shopify provide user feedback in the form of reviews [23]. Developers rely on this feedback to make informed decisions and prioritize their actions [5]. In recent times, there has been a growing inclination towards examining user reviews to extract insightful knowledge about software products and recognize areas for improvement. Although previous studies have been made to identify problems and concerns through user reviews [6], our study focuses on analyzing reviews that are specific to software ecosystems as analysis of ecosystems remains a challenge in software ecosystems [7].

Several studies have identified various problems in SECOs, such as coordination problems [8], vendor lock-in [9], interoperability issues [10], and project management [11]. The challenges of SECO research include understanding the complex interactions and selection of various stakeholders [12], developing effective governance mechanisms [13], designing appropriate business models [1], and Requirement elicitation [24]. The use of Natural Language Processing (NLP) and user review mining has become a popular research topic in software engineering due to the increasing importance of user feedback in software development [14, 15]. This approach involves analyzing user reviews to extract useful information, such as feature requests, bug reports, and user opinions. Work similar to ours has been on identifying privacy themes from user feedback [16] and classifying advertisement-related reviews [17].

However, analyzing software ecosystem reviews is difficult due to multiple feedback channels and the complex interplay with third-party applications. It can be hard to distinguish if feedback is for a single partner application, multiple applications, or the core platform [18]. Platform providers must rely on partners to gather feedback and make it accessible. The distinction between the core product and partner apps might become unclear, making it challenging for end users to provide feedback and platforms to analyze feedback [19]. To further our understanding of end-user challenges and their mitigation strategies in SECOs, we ask the following research questions:

- RQ1: What are the different problems faced by end-users in software ecosystems?
- RQ2: How has the amount of end user feedback in software ecosystems changed over time?
- RQ3: Are there recommended strategies for mitigating end-user challenges in software ecosystems?

1.1 Research Contributions

Our study provides several contributions. First, we introduce a method for researchers to work with user feedback in SECOs and distinguish SECO-related reviews. Then, we shed light on six areas of end-user concerns in software ecosystems and provide an array of discussion topics and feedback for each area. Additionally, we also reveal how SECO-related feedback has grown over time which shows the increasing need for studies in this space. Finally, we provide recommendations for developers and owners of software platforms to address and try to prevent these problems from occurring. The study’s two-part design enhances understanding of end-user concerns and industrial perspectives on software ecosystems, guiding platform design for better ecosystem management and sustainability through key roles keystones play in a platform’s success [1, 3, 20].

2 Methodology

We used a mixed-method study as summarized and illustrated in Fig. 1

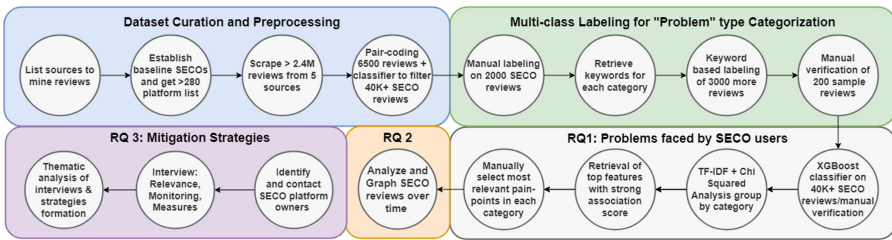


Fig. 1. Research Design Summary

2.1 SECO Platforms and Dataset Curation

First, we identified 15 popular SECO platforms, based on their characteristics such as integration, innovation, interoperability, marketplace, software as a service (SaaS), and integration platform as a Service (iPaaS) that define a SECO [1–3, 21] in addition to the well-defined classification of software ecosystems [4] as software platforms, service platforms, software standards. We further expand on the discussed “service platform” by categorizing them according to service sectors by selecting one or two platforms for each sector that serves as a baseline to retrieve similar platforms. We picked e-commerce platforms (*Shopify* and *WooCommerce*), CRM tools (*HubSpot*, *ZenDesk*, and *MailChimp*), Software as a Service (SaaS) (*SalesForce* and *Xero*), Communications Platforms (*Slack* and *Teams*), Payment Integration software (*Square Up*), Integration Platform as a Service (iPaaS) solutions (*Zapier*), development platforms (*Wix* and *WordPress*), and Human Resources Integration Platforms (*Bamboo HR*).

Table 1. User Feedback Collection

Source	Reviews Collected	SECO Reviews
TrustPilot	100,666	4,146
Google Play	1,396,059	17,089
App Store	159,595	1,778
Shopify Store	797,967	16,250
Other	998	998
Total	2,455,285	40,261

We retrieved applications from mobile application stores (Google Play and App Store) with search queries (*regex* = “software” + “as a service/platform/ecosystems/integration”) and by retrieving platforms “similar” to the identified 15 baseline platforms using Python libraries mentioned below. A total of 283 platforms were identified, but only **139** of them were used for analysis based on having SECO-relevant reviews (and which we discuss next). We used sources shown in Table 1 to collect user feedback from where we scraped **2,455,285** user reviews. The reviews were scraped using manual web scraping on TrustPilot, the *google-play-scraper*¹, and *app-store-scraper*² libraries in *Python*³ for respective Google and Apple app stores, *Kaggle*⁴ for Shopify store reviews, and directly from organizations. We combined all of it to form a single dataset with attributes ‘source’, ‘platform’, ‘review content’, ‘review date’, and ‘developer response’.

2.2 Identifying SECO-Related Reviews

To manually determine if a review is a SECO-related review, reviews were read in detail to understand the context of the user comments, employed pair coding and Cohen Kappa’s coefficient [22] in the process. The classification was further refined by utilizing SECO-related keywords such as “platforms,” “integration,” “API,” “ecosystems,” “plugins,” and “sync.” These keywords were instrumental in distinguishing SECO reviews from non-SECO reviews and were manually validated based on contextual understanding. For instance, reviews containing contextual clues such as integration issues, third-party app names, and plugin names were classified as SECO-related. Conversely, reviews that lacked explicit SECO-related terminology, such as those discussing poor app performance or usability issues, were classified as non-SECO reviews. Some reviews like “*the platform constantly crashes on my older iPhone..*” that at first appeared to be a SECO-related review, were classified irrelevant as well, as they do not provide specific challenge regarding use of the platform, rather a generic comment about compatibility.

¹ <https://github.com/JoMingyu/google-play-scraper>.

² <https://github.com/cowboy-bebug/app-store-scraper>.

³ <https://www.python.org/>.

⁴ <https://www.kaggle.com/>.

We began by creating a subset of 500 random reviews, ensuring an equal distribution of reviews corresponding to each rating scale, ranging from 1 to 5. A second-coder of the dataset labeled the identical 500 reviews with an author over 5 iterations of 100 reviews each, yielding an incremental agreement score, saturating at 0.81, indicating high agreement levels. Having built a shared understanding of what a “SECO-related review” is, we split 6000 random reviews (1200 reviews from rating 1–5 each). Upon combining the initial 500 reviews and the 6000 labeled reviews, a total of 848 SECO-related reviews were identified. Reviews like “*Nothing but issues with this platform. You change a setting and it doesnt work on *third-party app name*, fix it on *plugin name* and the platform changes it back!! Terrible Customer service dont help much, just tell you to speak to *platform name*! Who say its an integration issue. Wasted two days trying to integrate this and would have been quicker doing it all manually!*” were marked as a SECO review whereas reviews like “*Its a very useless app. It cannot run in normal internet speed. It’s a lot of confusion to use this app. It buffers a lot while attending class*” were marked as not relevant.

We then trained an XGBoost classifier [25] using the labeled 6500 reviews with a standard 80:20 proportion of train-test split for training the model. The model was trained with 0.97 accuracy, 0.99 precision, and 0.80 recall, and 0.89 F1-score, indicating high accuracy and reliability [26]. Having applied the 2.4 million reviews on this classifier, we were left with **40,261** reviews related to SECO from **139** platforms. Table 1 shows a breakdown of reviews retained from all the sources.

2.3 Manual Multi-class Labeling

On the 40,261 SECO-related reviews, we selected a balanced dataset (rating) of 2000 SECO-related reviews for manual labeling and further labeled 3000 more. We listed 6 common SECO issue themes and performed single-label, multi-class, manual classification following a well-practiced card-sorting technique [27]. Relevant keywords were created by observing term frequencies using TF-IDF Vectorizer [28] and manual observation. Categories and their keywords included: **Integration**: *integration, API, plugin, sync*; **Customer Support**: *customer, support, representative, speak*; **Design & Complexity**: *interface, confusing, easy, hard, design, customization*; **Privacy & Security**: *privacy, security, beware, fake, scam, login, authentication, password*; **Cost & Pricing**: *price, cost, refund, expensive, charge, buy, payment, credit, card, merchant, money*; **Performance & Compatibility**: *device, phone, slow, responsive, frequent, audio, video, crash, desktop, web, mobile, quality*. We used these keywords to label 3000 more reviews. A review belongs to a class with high confidence when at least 2 of the keywords were present in the review. If none two matched, at least one keyword need to be matched. If none of the keywords matched, they were simply classified as ‘Other’. We manually verified 200 randomized reviews and observed all of them accurately represented SECO-related concerns without any major overlapping of categories when filtered with at least 2 matching keywords.

2.4 SECO Challenges Classifier and Analysis Method

We used XGBoost⁵ as the primary classification model to classify reviews based on different categories. The dataset of 5000 train-test training reviews was pre-processed using well-used and known NLTK toolkit features⁶. We performed a training-test split with a frequently used ratio of 80:20. We used precision, recall, and F1-score as evaluation metrics to measure the performance [26] of the model in different categories. The XGBoost model achieved an accuracy of 0.93, with a macro average precision of 0.92, recall of 0.89, and F1-score of 0.90 as shown in Table 2, which indicates that the model was able to classify the reviews into different categories with very high accuracy. To validate the performance of the model, we manually verified a sample of 50 reviews from each category, which resulted in an accuracy of 91 percent. We compared the XGBoost model's performance with similar classification models. The XGBoost model outperformed with an accuracy of 0.93, while Linear SVC and Random Forest achieved an accuracy of 0.84 and 0.82, respectively. The methodology demonstrates the effectiveness of using XGBoost for classifying reviews into different categories.

Table 2. Classification Report

Label ID	Precision	Recall	F1-score
0	1.00	0.97	0.99
1	0.99	0.97	0.97
2	0.97	0.93	0.95
3	0.94	0.79	0.86
4	0.82	0.82	0.82
5	0.88	0.74	0.80
6	0.85	1.00	0.92
Accuracy			0.93
Macro Average	0.92	0.89	0.90
Weighted Average	0.93	0.93	0.93

We implemented the classifier on the 40,261 software ecosystem reviews. We identified the most relevant and frequently occurring terms (also referred to as features) using a set of *negative* reviews for each category. The set of negative reviews belonging to each category is kept using Vader Sentiment⁷ with a negativity score of over 0.4. The features present in those reviews are extracted using TF-IDF. In Eq. 1, t is a term (word), d is a document, D is

⁵ <https://github.com/dmlc/xgboost>.

⁶ <https://www.nltk.org/>.

⁷ <https://github.com/cjhutto/vaderSentiment>.

the corpus (collection of documents), 'tf' is the term frequency, and 'idf' is the inverse document frequency [28].

$$\text{tf-idf}(t, d, D) = \text{tf}(t, d) \cdot \text{idf}(t, D) \quad (1)$$

The reviews were preprocessed to remove non-English words, stop words, and tokenize them. We then performed Chi Squared analysis to measure the association between each feature and its' corresponding label. The chi-Squared analysis is a popular method not only for hypothesis validation but also useful for feature selection and computing association between features and their labels [29]. It can be implemented using the formula in 2 where χ^2 is the chi-squared statistic, n is the number of categories, O_i is the observed frequency in category i , and E_i is the expected frequency in category i .

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (2)$$

2.5 Interviews

Having identified these challenges, we also conducted qualitative research through semi-structured interviews [30] to derive and articulate a set of mitigation strategies. Four platform executives were selected for the interviews based on their roles, positions, and platform profiling (anonymized as P1, P2, etc.) as shown in Table 3. The selection used purposive sampling [31]. The interviewees were asked questions about monitoring user feedback, ensuring seamless integration, recommended strategies for solving challenges, managing an evolving marketplace of vendors, and other questions relating to the findings from RQ1.

Table 3. Interviewee Profile

Title	Company	Established	Size (employees)	Country
Chief Technology Officer	P1	2017	100-200	Canada
VP Engineering	P2	2013	50-100	Canada
Platform Ecosystem Advocate	P3	2006	8000-10000	USA
Chief Technology Officer	P4	2017	300-500	Nepal

The interviews were conducted following ethical principles, including informed consent, confidentiality, and privacy, as per university approved research ethics application. The data collected from the interviews were transcribed, sorted, and analyzed using a thematic analysis approach [32], which enabled us to identify and analyze the themes and patterns in the data related to how companies identify and address issues related to software ecosystems through user feedback.

3 Findings and Discussion

3.1 Distribution

Out of 40,261 reviews, ‘Integration’ has the highest proportion of software ecosystem reviews at 28.85% with a 4.26/5 median rating. ‘Customer Support’ is the second highest category at 17.67% with a 3.72/5 median rating, followed by ‘Design and Complexity’ at 8.35% with a 4.47 rating. ‘Privacy and Security’ have the lowest rating of 2.87/5 with 4% of the reviews, ‘Cost and Pricing’ has 6.74% with a 3.67/5 rating, and ‘Performance and Compatibility’ has the lowest proportion of reviews at 2.80% with a 3.78/5 median rating. SECO review not fitting into any of the six categories were classified as ‘Other’ with 31.58% of the reviews, leaving room for future work for introduction of additional categories.

3.2 RQ1: End-User Pain-Points in SECOS

In this section, we present the findings from reviews for all classified areas of SECO issues. In order to extract the pain-points (features), we performed the following set of operations: Let C be a set of reviews with respective category IDs, where review r_i has a sentiment score $s_i \in \text{positivescore}, \text{negativescore}$. Let $C = (l_i, R_i) \mid i = 1, 2, \dots, n, s_i = \text{negative score} > 0.50$ be the set of negative reviews. Let $L = l_1, l_2, \dots, l_n$ be the set of categories present in C . Define $R_i = r_j \mid r_j \in R_i$ and $s_j = \text{negative}$ as the set of negative reviews belonging to category l_i . Define TF-IDF $_c : R_c \rightarrow F$, where $F = (r, f) \mid r \in R, f \in W$ is the set of review features for all reviews in C . Let $F'_l = f \mid (r, f) \in \text{TF-IDF}_c(R), r \in R_l$ be the set of features present in reviews of category l . Let $\chi^2(f, l)$ be a statistical measure of association between feature f and category l . Then, the set of categories and their top 100 features with a $\chi^2(f, l)$ is given by: $(\text{Labels}, (\text{feature}, \text{score}))[1, 100] = (l, F'_l, \chi^2(f, l)) \mid l \in C, f \in F'_l, \chi^2(f, l)$.

Integration. The first category of pain points in software ecosystems is related to integration, with the most common issues being problems with integration and a “lack” of integration altogether. These are followed by “cross-platform issues”, “API errors”, and “API key” problems. Users are frustrated with the difficulty of integrating different software components and systems, which leads to inefficiencies and lost productivity. One of the most common integration complaints is regarding “Facebook API” errors. Similarly, integration errors with “Google API” caused issues with SEO and other critical aspects of online business. Another common integration issue mentioned in the data is the lack of “PayPal integration”. “Mailchimp integration” and “Outlook Integration” are other common issues that cause problems with email marketing campaigns. Several of the pain points in this category are related to specific platforms, such as “Android integration”. The pain points related to integration in software ecosystems can have significant impacts on software architecture [33].

Customer Support. The second category of pain points in software ecosystems is related to customer support extracted from SECO-related reviews. The top pain point in this category is “worst customer service”, followed by “impossible to reach”, “service joke/rude”, and “speak English” indicating significant dissatisfaction among users with the customer service provided by the software ecosystem. Other pain points include difficulty reaching customer support and poor quality of service. Customers seem to prefer speaking to “real humans” over “chat. Poor customer service could result in lost customers and damage to the organization’s reputation. Platforms may need to invest in better support channels to ensure that users and third-party developers have access to the help they need. Overall, the problems identified suggest that users have a variety of dissatisfaction with the customer support provided by the platforms.

Design and Complexity. In our study, the most frequent pain point in the user experience category is around the topic of “bad user interface”. This can be evaluated in several ways from previously established theories [34] and our own findings such as problems in “sorting” and “ads”. Some of the other topics provide more specific examples of what users find challenging about the software interface. For example, the “mobile app interface” topic showed that users have difficulty with software that is primarily mobile-based. The “web interface” related reviews mentioned that users find web-based software challenging to navigate. Additionally, “interface slow” and “lags” indicate that users have problems with the performance of the software. Issues such as “desktop interface” and “other app easy” indicate that users have trouble with desktop-based software and that they may compare it unfavorably to other, more user-friendly applications. The topics in this category suggest that users find software with bad or confusing user interfaces frustrating and difficult to use, which can lead to decreased productivity, innovation, and satisfaction with the software.

Privacy and Security. Privacy and security are critical concerns for most software users, especially in the e-commerce platform realm [35]. Users are often hesitant to trust a platform with their personal and sensitive information [36], and the reviews in this category reflect that. The features discussed in this category include “possible scams”, “fake apps”, and “fake reviews”, all of which suggest that users are worried about the legitimacy of the platform and the third-party apps they are using. Some important pain points in this category were “impossible login” and “keeps asking for passwords”, indicating that users are struggling to access their accounts. An interesting issue topic identified is “data mining”, showing that users are concerned about how platforms are mining their personal data. Other pain points in this category relate to user authentication and security measures. The issue topic of marketplace scammers suggests that users are worried about fraudulent third-party marketplace sellers on the platform. Platforms that can address these concerns and implement robust security measures by clearly stating policies, increased lucidity, and readability are likely to have happier and more trusting users [37].

Cost and Pricing. Pricing is an important characteristic of ecosystem marketplace [38, 39]. This category focuses on the cost and pricing structures of SECOs. The main pain points raised by users were related to “losing money”, “issues with credit card payments”, and “expensive fees”. The reasons for this were “unexpected charges”, “hidden fees”, and ineffective “refund policies”. The pain point “credit card” had a significant association score, indicating that users had issues with their card payments. The pain point “waste money” indicated that users felt that they were spending money on a product that was not worth the cost. Other pain points related to cost and pricing include “refund impossible”, “prices expensive”, “fees expensive”, and “charged accounts”. These raised issues suggested that users lost the company’s trust and were dissatisfied with the pricing and fees associated with the platforms and their services and that they had difficulty obtaining refunds or finding affordable alternatives.

Performance and Compatibility. Though companies choose cross-platform development more and more over native development [40] the most significant pain points in this final category seem to be “web interface” and “device version”, followed closely by the topic “multiple devices” and “loss connection”. These pain points suggest that users are experiencing sync and connectivity issues across web, desktop, and mobile versions of the platform. Another common topic in this category is “mobile website”, suggesting that users are having difficulty accessing and using the software ecosystem on their mobile devices. The pain point “loss data work” suggested that users are experiencing data loss or data corruption while using the software ecosystem. Other pain points in this category included “video audio quality”, “lost quality”, “iPhone iPad issue”, “don’t trust app”, “phone horrible”, “buggy slow”, “app crashes constantly”, “web version”, “loss clients”, “phone laptop”, “sort problem”, and “messed website”. These pain points suggest that users are experiencing issues with the overall functionality and reliability of the software ecosystem, causing them to lose trust in platforms, and even instances of businesses losing clients.

3.3 RQ2: Growth in SECO Feedback Over-Time

We analyzed the change in SECO-related review numbers over time by mapping the reviews from January 2013 to December 2022. We grouped the reviews by month and counted the number of reviews in each month. We calculated the median count for all categories. Reviews from before 2013 and from 2023 were discarded due to their insignificance in number.

We can observe from Fig. 2 that there has been a significant rise in software ecosystem reviews in the last decade, with the reviews regarding SECOs starting to grow significantly from 2016 onwards. The number of SECO reviews increased from 51 in 2013 to 4,610 in 2022, with the highest growth occurring between 2016 and 2020. In 2020, the growth rate went to a 130.08 percent increase from 2019, but it declined in 2022 with a -26.75 percent growth rate compared to the previous year. The average growth rate from 2018 to 2022 was **258.11** percent.

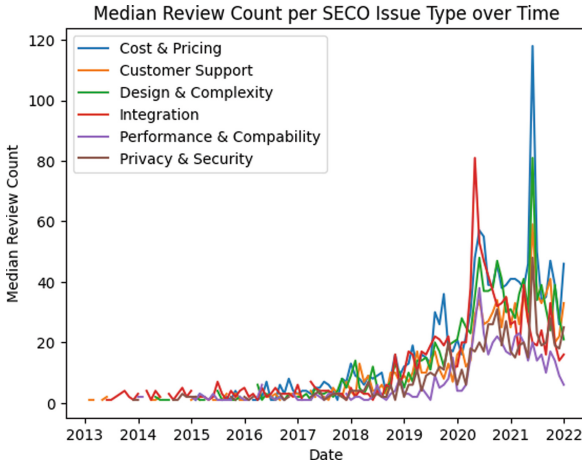


Fig. 2. Change in SECO reviews over time

From our interviews, we confirmed that platform organizations faced an increasing demand for integration tools and customer support during the COVID-19 pandemic⁸.

3.4 RQ3: Mitigation Strategies for Platforms

Here, we present our interview findings with platform owners in the form of recommendations, who also fully validated the challenges discussed earlier.

API First Approach. Application Programming Interface (API) first development is a strategy that focuses on building the API first before allowing third-party developers to make an integration request. This prevents organizations from having to implement one-off integration specific to the developer request. For example, the VP of Engineering from P2 said “...*small startups have an API first mentality. It’s in the DNA of the company that they’re building an API so that they don’t run into one-off issues.*”, which potentially addresses the most talked about API-related end-user concerns such as “lacks integration”.

User and Developer Communities. Mitigating customer support and other end-user problems in a software ecosystem requires actively engaging the user community, supporting developers, continuously improving the platform, and fostering collaboration and partnerships. These strategies help address issues, enhance the user experience, and align with evolving integration requirements, as quoted by P4’s CTO, “...*an ecosystem doesn’t thrive if there’s no community for all the stakeholders.*”.

⁸ <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.

Third-Party App Control. Platform owners should mitigate security and financial risks and issues in their ecosystem by implementing a strict vetting process, continuously monitoring and auditing third-party apps, incentivizing safe and high-quality apps through pricing strategies, and providing developer support and resources. Platform P3’s advocate says *“If somebody had essentially abandoned all supported their app and they would be removed from our marketplace”* which ensures compliance and monitoring in the marketplace.

Feedback-Driven Approach. In order to effectively mitigate design, complexity, and performance issues, adopting a feedback-driven approach is a valuable strategy for platform owners. As mentioned by the CTO of P4 *“We monitor user interactions within the apps. We get notices of, like rage clicks, things like that, where they go.”*, implementing tracking tools, actively soliciting and carefully prioritizing feedback, incorporating user and developer input into the development process, and maintaining transparent communication channels are advisable.

Cross-Platform Development. Platform owners should prioritize cross-platform development and utilize progressive web apps (PWAs) to enhance the platform’s accessibility and provide a consistent user experience across different devices. To quote P1’s CTO, *“We would consider like a cross-platform Progressive Web App To make everything work with mobile devices across the board”*, extending the platform’s reach and maintaining competitiveness through cross-platform development, platform owners can attract a wider audience and mitigate platform-specific user issues.

Documentation and Guidelines. Platform owners should prioritize comprehensive documentation, accessibility, quality and security guidelines, and developer support in optimizing the utilization of the platform’s API. By providing clear instructions, easy access, and assistance to developers, platform owners can foster a collaborative and productive developer community, resulting in high-quality integration and improved platform success, as P3’s advocate said, *“It starts with having really clear ATP documentation. I think having that publicly available, they start first ideating about the process.”*

User Data Management. By providing transparent policies, establishing efficient incident response processes, prioritizing user privacy, and adhering to relevant regulations, platform owners can foster trust, protect user information, and mitigate potential risks associated with data breaches or non-compliance. For example, P1’s CTO said, *“We don’t hold the client information in our databases for any longer than, you know, The lifetime of an order which is the lifecycle of the data.”*, and P2’s VP of engineering mentioned *“Good user data management practice such as streamlined SSO authentication is a good practice to resolve integration as well as privacy issues”*, meaning platform owners must ensure that third-party applications delete user data when it is no longer needed, and secure authentication practices must be implemented.

4 Implications

This study represents a first large-scale investigation of end-user challenges in software ecosystems. We presented a method for identifying user feedback that distinguishes SECO-related reviews from general reviews by using methods explained in Sect. 2.2. We also identified that integration issues, customer support, the complexity of design and user interface, issues with privacy and security, pricing issues, and platform compatibility are problem areas in software ecosystems, as well as a set of recommendations to mitigate these challenges. This study has significant implications for SECO researchers, highlighting unexplored end-user challenges and the lack of prior research. The temporal growth of SECO-related reviews, particularly during the COVID-19 pandemic, underscores the dynamic nature of SECOs. The study's recommendations offer actionable guidance for both researchers and industry stakeholders.

5 Threats to Validity

The study's results may be influenced by the varied quality and accuracy of data from different sources and limited interviews. The user feedback, mainly from mobile app reviews, may not fully represent all users across various software platforms. The data, although extensive, was selectively scraped from certain platforms, potentially limiting its applicability to diverse software ecosystems, especially open-source software. The identification of software ecosystem-related issues was crucial to the analysis which is a potential threat to the construct validity. However, the pair-coding approach with inter-rater agreement was the most ideal way of initially classifying what a SECO review is. Also, manually investigating the results of the automated classification to ensure accuracy alongside an optimal evaluation results of the classifier.

6 Conclusion and Future Work

This study provides a valuable contribution to the existing knowledge of end-user concerns and the industrial perspective on software ecosystems. By identifying key issues and providing recommendations in several aspects of a SECO platform, our findings can guide platforms in designing and fostering better ecosystems. The methods and techniques used in this study can serve as methodological guidance for future research in this space.

Future work could expand the scope of the study to include more ecosystem platforms and user reviews. The two machine learning classifiers could be further refined to improve its accuracy in first identifying what kind of feedback is a SECO-related feedback, and secondly in categorizing SECO reviews according to the proposed problem areas. Additional problem categories could be identified and analyzed. The effectiveness of the mitigation strategies suggested could be evaluated through implementation and user feedback. Longitudinal studies could be conducted to track the changes in user challenges and developer responses over time.

References

1. Jansen, S.: A focus area maturity model for software ecosystem governance. *Inf. Softw. Technol.* **118**, 106219 (2020)
2. Manikas, K., Hansen, K.M.: Software ecosystems - a systematic literature review. *JSS* **86**(5), 1294–1306 (2013)
3. van den Berk, I., Jansen, S., Luinenburg, L.: Software ecosystems: a software ecosystem strategy assessment model. In: *Proceedings of the 4th ECSA: Companion Volume*, pp. 127–134. ACM (2010)
4. Jansen, S., Cusumano, M.A.: Defining software ecosystems: a survey of software platforms and business network governance. In: *Software Ecosystems*, pp. 13–28. Edward Elgar Publishing (2013)
5. Nguyen, D.C., Derr, E., Backes, M., Bugiel, S.: Short text, large effect: measuring the impact of user reviews on android app security & privacy. In: *2019 IEEE Symposium on SP*, pp. 555–569 (2019)
6. Genc-Nayebi, N., Abran, A.: A systematic literature review: opinion mining studies from mobile app store user reviews. *JSS* **125**, 207–219 (2017)
7. Serebrenik, A., Mens, T.: Challenges in software ecosystems research. In: *Proceedings of the 2015 ECSA Workshops*, Art. 40. ACM, New York, NY, USA (2015)
8. Figalíst, I., Elsner, C., Bosch, J., Olsson, H.H.: Scaling agile beyond organizational boundaries: coordination challenges in software ecosystems. In: Kruchten, P., Fraser, S., Coallier, F. (eds.) *XP 2019. LNBIP*, vol. 355, pp. 189–206. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19034-7_12
9. Rani, A., Mishra, D., Omerovic, A.: Rani, A., Mishra, D., Omerovic, A.: Multi-vendor software ecosystem: challenges from company’ perspective. In: Rocha, A., Adeli, H., Dzemyda, G., Moreira, F. (eds.) *Information Systems and Technologies. WorldCIST 2022. LNNS*, vol. 470, pp. pp. 382–393. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-04829-6_34
10. Andreo, S., Bosch, J.: API management challenges in ecosystems. In: Hyrynsalmi, S., Suoranta, M., Nguyen-Duc, A., Tyrväinen, P., Abrahamsson, P. (eds.) *ICSOB 2019. LNBIP*, vol. 370, pp. 86–93. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33742-1_8
11. Ghimire, B.: Assessing the Effectiveness of Project Management on Software Engineering. Master’s thesis, Collegium Humanum (2021)
12. Knauss, E., Yussuf, A., Blincoe, K., Damian, D., Knauss, A.: Continuous clarification and emergent requirements flows in open-commercial software ecosystems. *Require. Eng.* **23**, 97–117. Springer (2018). <https://doi.org/10.1007/s00766-016-0259-1>
13. Baars, A., Jansen, S.: A framework for software ecosystem governance. In: Cusumano, M.A., Iyer, B., Venkatraman, N. (eds.) *ICSOB 2012. LNBIP*, vol. 114, pp. 168–180. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30746-1_14
14. Madhurika, B., Malleswari, D.N.: A systematic review on artificial intelligence-based opinion mining models. In: *2022 International Conference on Edge Computing and Applications (ICECAA)*, pp. 252–257. IEEE (2022)
15. Solangi, Y.A., Solangi, Z.A., Aarain, S., Abro, A., Mallah, G.A., Shah, A.: Review on natural language processing (NLP) and its toolkits for opinion mining and sentiment analysis. In: *2018 IEEE 5th ICETAS*, pp. 1–4. IEEE (2018)
16. Nema, P., Anthonysamy, P., Taft, N., Peddinti, S.T.: Analyzing user perspectives on mobile app privacy at scale. In: *Proceedings of the 44th ICSE*, pp. 112–124. ACM (2022)

17. Gao, C., Zeng, J., Lo, D., Xia, X., King, I., Lyu, M.R.: Understanding in-app advertising issues based on large scale app review analysis. *IST* **142**, 106741 (2022)
18. Foundjem, A.: Cross-distribution feedback in software ecosystems. In: Proceedings of the IEEE/ACM 42nd ICSE Workshops, pp. 723–724. ACM (2020)
19. Johnson, D., Tizard, J., Damian, D., Blincoe, K., Clear, T.: Open crowdRE challenges in software ecosystems. In: 2020 4th CrowdRE, pp. 1–4. IEEE (2020)
20. Belo, Í., Alves, C.: How to create a software ecosystem? A partnership meta-model and strategic patterns. *Information* **12**(6), 240 (2021)
21. Jansen, S., Cusumano, M.A., Brinkkemper, S.: *Software Ecosystems: Analyzing And Managing Business Networks in the Software Industry*. Edward Elgar Publishing (2013)
22. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. *Biometrics*, 159–174. JSTOR (1977)
23. Jansen, S., Bloemendal, E.: Defining app stores: the role of curated marketplaces in software ecosystems. In: Herzwurm, G., Margaria, T. (eds.) ICSOB 2013. LNBP, vol. 150, pp. 195–206. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39336-5_19
24. Damian, D., Linåker, J., Johnson, D., Clear, T., Blincoe, K.: Challenges and strategies for managing requirements selection in software ecosystems. *IEEE Softw.* **38**(6), 76–87 (2021)
25. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD, pp. 785–794 (2016)
26. McIlroy, S., Ali, N., Khalid, H., Hassan, A.E.: Analyzing and automatically labelling the types of user issues that are raised in mobile app reviews. *Empir. Softw. Eng.* **21**, 1067–1106 (2016)
27. Spencer, D.: *Card Sorting: Designing Usable Categories*. Rosenfeld Media, New York (2009)
28. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *JMLR.* **12**, 2825–2830 (2011)
29. Schütze, H., Manning, C.D., Raghavan, P.: *Introduction to Information Retrieval*, vol. 39. Cambridge University Press, Cambridge (2008)
30. Longhurst, R.: Semi-structured interviews and focus groups. In: *Key Methods in Geography*, vol. 3, no. 2, pp. 143–156 (2003)
31. Tongco, M.D.C.: Purposive sampling as a tool for informant selection. In: *Ethnobotany Research and Applications*, vol. 5 (2007)
32. Clarke, V., Braun, V., Hayfield, N.: Thematic analysis. In: *Qualitative psychology: A practical guide to research methods*, vol. 3, pp. 222–248 (2015)
33. Bosch, J.: Architecture challenges for software ecosystems. In: Proceedings of the 4th ECSA: Companion, pp. 93–95 (2010)
34. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI CHI, pp. 249–256 (1990)
35. Udo, G.J.: Privacy and security concerns as major barriers for e-commerce: a survey study. *Inf. Manage. Comput. Secur.* **9**, 165–174 (2001)
36. Aïmeur, E., Lawani, O., Dalkir, K.: When changing the look of privacy policies affects user trust: an experimental study. In: CHB, vol. 58, pp. 368–379 (2016)
37. Ermakova, T., Baumann, A., Fabian, B., Krasnova, H.: Privacy policies and users' trust: does readability matter?. In: AMCIS (2014)
38. Burkard, C., Draisbach, T., Widjaja, T., Buxmann, P., et al.: Software ecosystems: vendor-sided characteristics of online marketplaces. In: GI-Jahrestagung, p. 421 (2011)

39. Jansen, S., Brinkkemper, S., Finkelstein, A.: Business network management as a survival strategy. In: Software Ecosystems, pp. 29–42 (2013)
40. Xanthopoulos, S., Xinogalos, S.: A comparative analysis of cross-platform development approaches for mobile applications. In: Proceedings of the 6th BCI, pp. 213–220 (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





A Survey on Perceptions of Data Sharing in the Norwegian Public Sector

Leif Z. Knutsen¹ , Bertha Ngereja¹ , Ingebjørg Flaata Bjaaland¹ ,
Jo E. Hannay¹ , and Sinan S. Tanilkan² 

¹ Simula Metropolitan Center for Digital Engineering, Pb. 4 St. Olavs plass, 0130 Oslo, Norway
{leif,bertha,ingebjorg,johannay}@simula.no

² Norwegian Computing Center, Pb. 114 Blindern, 0314 Oslo, Norway
sinan@nr.no

Abstract. Sharing data among public institutions is essential for reaping the benefits of data-driven capabilities. Literature to date has identified several types of benefits that are likely to accrue to a wide range of sectors, as well as challenges and obstacles to implementing data-sharing solutions. We sought to identify perceptions of possible benefits, likely challenges, and the likelihood of overcoming them in the Norwegian public sector. Our survey of IT practitioners interested in the subject suggests that optimism about data sharing is high, concerns about a wide range of challenges are also high, and confidence in public institutions is tenuous. Responses also suggest that divisional management may be critical in implementing data sharing solutions. The pattern of responses suggests uncertainty consistent with low maturity in the field. We posit that data sharing among public institutions is part of a broader set of capabilities needed for public service innovation across institutions.

Keywords: Data Sharing · Public Sector · Survey · Digitalization

1 Introduction

Digital innovation in the public sector depends on the effective and responsible use of data that public institutions collect, use, generate, and share. There is considerable optimism about the potential benefits of data-oriented capabilities. For example, *open data* – making specific data publicly accessible, reliable, and understandable [25] – is associated with better use of data and better services [28]. *Big data* has several connotations [17] but refers broadly to the ability to perform analyses and generate insights from large, often exhaustive datasets. It has been identified as a driver of public-sector innovation [26,43]. Being *data driven* is seen as a strategic capability [32] and as an element for restructuring the public sector [24].

As capabilities in the public sector [35], open and big data highlight the need for governments to gather and collate data from disparate sources. Thus, the ability to *share data* is a prerequisite for both big and open data and other data-oriented capabilities in the public sector. But also as a fundamental capability in itself, *data sharing* – the ability to share data among public and private institutions to improve the value and

quality of services and to increase the scope of data available to decision-makers – creates opportunities for improving government services [11, 13, 32, 36].

Public institutions have the legal authority to collect a wide range of data sets, but they also have the legal responsibility to safeguard them against abuse, disclosure, or damage. In addition to comprehensive legislation that restricts the use of data, such as the General Data Protection Regulation (GDPR) in the European Union (EU), several policy issues have been raised [9, 10, 47]. The practice of data sharing, i.e., that public institutions exchange data with each other, with private sector parties, and even across national boundaries, attracts concern. For example, GDPR allows organizations only to use data for disclosed purposes. Notwithstanding these constraints, institutions such as the EU see data sharing as an important part of improving government services [49], leading to a tension between realizing the full range of benefits from data sharing on the one hand and protecting citizens' rights on the other [21]. Governments also face obstacles in realizing the benefits of data sharing, such as restrictive legislation and policies, bureaucratic boundaries, diverse procedures in institutions, lack of trust, lack of resources, technical issues, and more [29, 50].

Norway's public sector is based on a unitary form of government with responsibility for services devolved to local governments and regional organizations. Public institutions maintain registers for individuals, companies, property, and more. Some data is shared among both public and private institutions for specific purposes, for example generating tax documents. There are calls for further data sharing, for example, health data among general practitioners and hospitals.

Moreover, a group of IT executives in the Norwegian public sector (Skate – Management and coordination of services in e-government) has taken several initiatives to capitalize better on authoritative data registers by sharing data among public institutions, both “vertically” between national and local authorities, and “horizontally” between public institutions at the same level.¹ The prospect of ensuring better health outcomes has motivated significant efforts to ensure sharing of health data [15].² Articles in the public press express frustration about the lack of progress in this area [16].

It falls to IT practitioners to realize the benefits of data sharing and overcome barriers. The motivation for the present study is to understand better IT practitioners' level of interest in this topic and their perceptions of both the promises and the difficulties of data sharing.

2 Background

In the literature, characteristics of data sharing for public services have been described in terms of areas in which data sharing applies, including anticipatory government, service design and delivery, and performance management [32]; in terms of at what level data is shared: technical, organizational and political [13, 36], and in terms of the types of benefit data sharing might yield, such as innovation, transparency, and efficiency [11].

¹ <https://www.digdir.no/skate/rad-til-regjeringens-digitaliseringsarbeid/3034>.

² <https://www.digdir.no/digitaliseringsradet/direktoratet-e-helse-helsedataprogrammet-2018/1998>.

Authors have applied different paradigms for categorizing obstacles and challenges to data sharing including impediments related to control, management, lacking agreement on goals, long goals, and lack of funding [36]; challenges related to obtaining useful data, data sharing, interoperability, discoverability, human and technical capacities, and legitimacy and public trust [32], public manager uncertainty about big data [22], digital champions' perceptions of barriers [48]; issues that may be cultural and political, technical, related to privacy and security, and efficient data management [11].

We have, however, yet to find research on the perceptions that IT practitioners might have about issues concerning data sharing. Consequently, we seek to build an understanding of IT practitioners' level of interest in the topic, their perceptions of benefits, their perception of challenges and hindrances, their perception of the benefits of data sharing certain segments of the public sector, their perception on funding data sharing and finally, their confidence in the public sector's ability to realize opportunities/benefits and overcome challenges/obstacles. We briefly recount relevant literature on each of these themes.

2.1 Benefits of Sharing Data

Articulating, measuring, and managing benefits in the public sector involves challenges [40]. One issue is that benefits may accrue to more than one actor and in some cases do not benefit the sponsoring institution at all. Several schemata have been proposed for disaggregating potential benefits of data sharing. To capture perceptions, we chose and adapted classifications that, in our experience, were relevant to the public sector. As a starting point, Christodoulou et al. [11] provided three areas for which data sharing can provide benefits (innovation, transparency, and efficiency), and we added elements from other research; i.e., case processing, decision-making, [6], data collection [2], error correction [42], and productivity [13]. These benefits areas are summarized in the upper-left portion of Table 1.

2.2 Challenges and Hindrances to Sharing Data

If sharper clarity on the benefits of sharing data drives more and better-targeted data sharing solutions, a clearer understanding of challenges should prepare practitioners and reduce the likelihood of delays and other problems. The literature has surfaced different challenges and hindrances related to internal capabilities, lack of shared standards that enable sharing, and other external limitations, especially regulatory and legal. From the literature, we derived the following specific types of challenges and hindrances: leadership support and legal/regulatory issues [4, 38], shared technical infrastructure [19, 27], strategic approaches [3, 14], technical standards [13], common semantics [46], short-term versus long-term goals [29], and technical competence [6]. These are summarized in the upper-right portion of Table 1.

2.3 Data Sharing in Different Public Sector Segments

The Organisation for Economic Co-operation and Development (OECD) uses the Classification of the Functions of Government (COFOG) [8, 31], which we found to be generally applicable but too broad at its highest (divisional) level and too granular at lower

Table 1. Concepts of data sharing

<i>Benefits areas</i>	<i>Challenges and Hindrances</i>
Innovation [11]	Leadership support [4,38]
Transparency [11]	Legal and regulatory issues [4,38]
Efficiency [11]	Shared technical infrastructure [19,27]
Case processing [6]	Strategic approaches [3,14]
Decision making [6]	Technical standards [13]
Data collection [2]	Common semantics [46]
Reducing errors [42]	Short-term vs long-term goals [29]
Productivity [13]	Technical competence [6]
<i>Public-sector segments [31]</i>	<i>Funding</i>
Healthcare	Earmarked funding [45]
Welfare	In competition with cross-segment funding [45]
Defence and National security	Internal budgeting in each organization [45]
Services for Businesses	Philanthropic donations [34,45]
Agriculture	Contributions from collaborating organizations [20,33]
Police, Customs, etc	
School and Education	
Higher Education	
Research	
Public Finance	
Children and adolescents	
Transportation and Infrastructure	
Environment and Sustainability	
Art and Culture	
Cross-sectorial	

levels. Based on a survey and analysis of IT activity and expenditures by government agencies we conducted in 2021 (currently unpublished), we elaborated the COFOG logic and created a classification intended to be more intuitive for IT professionals, summarized in the lower-left portion of Table 1.

2.4 Funding Data Sharing Initiatives in the Public Sector

Funding is an important factor for data sharing in the public sector [5,23,51]. Developing and implementing data sharing initiatives are costly in both tangible (people, money, equipment) and intangible aspects (data, information), while the benefits are often hidden and unclear, leading the government to opt spending their budget on other investments [7]. Nonetheless, the governmental ability and readiness to invest in the necessary digital innovations and its related costs are essential [5].

Public-sector policy frameworks for funding initiatives may well result in implications such as the lack of reliable and dedicated funding for the cross-boundary collaboration and cooperation that is necessary for sharing data [33,51]. Since data sharing initiatives in the public sector are initiated on an ad-hoc basis, they may only sometimes be prioritized against other initiatives considered as more critical [51]. Consequently, data sharing initiatives in the public sector, in general, are hindered by financial constraints [5,23,51]. In the following, we elicit relevant funding alternatives that we summarize in the lower-right portion of Table 1.

The traditional alternative is to allocate government budget through fixed-term stable funding [45], but this approach may not work well for digital innovations because it does not take into account the long-term funding requirements and the need for collaboration across organizations [5] and may require maintenance and further development. Funding plans should include the maintenance process and resources [45]. Alternatives to traditional fixed-term funding should be considered [45]. One flexible approach suggested is stable fixed-term funding with the flexibility to be provided annually as the initiative is developed [45].

In addition to constraints imposed by government budgeting and funding practices, data sharing initiatives in the public sector face funding challenges with approaches that are unstable over the time horizons of data-sharing solutions. Examples of these unstable approaches include (i) grants and funding programs [45], (ii) institutional funding [45], (iii) philanthropic donations from foundations [34,45], or (iv) external funding from strategic partnerships with other organizations [20,33]. The challenge with external funding is that data sharing may stop when the funding ends [20].

3 Research Questions

The manifold issues above on realizing benefits and overcoming obstacles, and our interest in better understanding IT practitioner perspectives leads us to formulate the following research questions:

- RQ1:** To what extent are IT practitioners interested in data sharing as a topic?
- RQ2:** To what extent do practitioners perceive that data sharing can contribute to the benefits areas of Table 1?
- RQ3:** To what extent do practitioners perceive that data sharing can create value in the public-sector segments of Table 1?
- RQ4:** To what extent do practitioners perceive that the challenges and hindrances of Table 1 impact good data sharing solutions?
- RQ5:** How appropriate do practitioners think that the funding alternatives of Table 1 are for data sharing?
- RQ6:** How much confidence do practitioners have in the public sector's ability to realize the potential value and overcome hindrances? If applicable, how confident are they about their own organization's abilities?

4 Methodology

We operationalized the concepts in the research questions in a manner intended to have relevance for the particular study setting of a seminar for Norwegian IT professionals.

4.1 Survey Design

Table 2. Survey questions

	<i>Survey questions</i>	<i>Answer options</i>
SQ1	How large is your interest in data-sharing (on three interest variables)?	11-point ordinal (no – large interest)
SQ2	How familiar are you with the possibilities and challenges associated with sharing data in the public sector (on three familiarity variables)?	11-point ordinal (low – very confident)
SQ3	How much do you think data sharing can contribute to improvement (in eight benefits areas)?	11-point ordinal (little – much)
SQ4	How useful do you think data sharing is for the following (15 segments) of the public sector?	11-point ordinal (not useful – useful)
SQ5	How much do you think the following (nine challenges) hinders good data sharing solutions?	11-point ordinal (little – much)
SQ6	How much confidence do you have in the public sector meeting the following (six requirements) for data sharing?	11-point ordinal (little – much)
SQ7	How suitable do you think the following (four mechanisms) are for funding data sharing among organizations over a five-year period?	11-point ordinal (poorly – well suited)
SQ8	How well do you think your organization succeeds in (two action variables)? ^a	11-point ordinal (poorly – well)

^a Asked only to those reporting to work in an organization where data sharing is relevant

We designed an online questionnaire starting with demographic questions about the respondents' organizational level of responsibility, functional area, and whether they worked in the public or private sectors; their personal interest in data sharing; and perceived knowledge about the topic at hand. Following this, the main part of the questionnaire contained sections based on the concepts summarized in Table 1. The survey questions directly relevant to answering the research questions are in Table 2. The full questionnaire design (in Norwegian and the English translation), the survey results and full analysis can be found at <https://osf.io/a53nx/>.

4.2 Survey Execution

We ran the survey in late August 2023 at a seminar titled “Sharing of Data among Actors – opportunities, limitations, and solutions”.

Forty-seven people attended the seminar in person, and 28 attended online, yielding $n_{\text{total}} = 66$ responses. Five provided demographic data only, leaving $n_{\text{included}} = 61$ responses answering SQ1–SQ8, which is the set of responses included in the analysis. Two respondents replied only to SQ1 and SQ2, and one replied to all questions until SQ7 (but not SQ8), leaving $n_{\text{complete}} = 58$ respondents who completed the entire survey. (Respondents were allowed to leave questions unanswered.) Among the n_{included} respondents, 4.0% worked in top management, 11.5% in divisional management, 50.8% as project or team leaders, 27.9% as specialists or experts and 4.9% in other work areas.

Respondents' area of daily work was: 36.1% technology, 34.4% development, 14.8% staff functions, 4.9% in the line organization, and 9.8% reported other.

Further, 32.8% were employed in the private sector (54.9% of these were allocated to an assignment for the public sector), and 67.2% were employed in the public sector, bringing the total of respondents whose daily work is in the public sector to 86%.

4.3 Survey Data Analysis

We present quartile boxplots for visual inspection of the results. We conducted ordinal comparisons between the variables in Table 1 with *Friedman's two-way analysis by ranks*, reporting omnibus tests across all variables and pairwise comparisons between pairs of variables. For each variable, we further conducted categorical comparisons between the organizational levels and also between the work domains with the *independent samples Kruskal-Wallis* test for three or more categories of data, reporting omnibus tests across all categories and pairwise tests between categories. These non-parametric tests are suitable because we cannot make assumptions about the distributions in the variables [30].

We accept a significance level of $\alpha = 0.05$; i.e., that a difference in our sample between variables or categories has a 5% (or lower) probability of falsely indicating a difference in the population. Here, we only report significant results due to space restrictions. All tests and descriptive statistics are generated using *IBM SPSS* (v.27).

We report effect size for the Kruskal-Wallis test using Cohen's d ,³ with the following rules of thumb: <0.1 (very small), $0.1 - <0.3$ (small), $0.3 - <0.5$ (medium) and $0.5 - <1.2$ (large) [12, 39]. For Friedman's tests, effect size estimates are calculated in terms of Kendall's W [44]. As Kendall's W has a different range from Cohen's d , different rules of thumb are needed to evaluate effect sizes for Kendall's W : $0.1 - <0.3$ (small), $0.3 - <0.5$ (medium) and ≥ 0.5 (large) [39]. These effect size measures only apply at the omnibus level [41]. Where applicable, we report the corresponding omnibus effect size as a proxy for effect sizes for pairwise comparisons.

5 Results

RQ1: IT Practitioners' Interest in Data Sharing. Figure 1 shows boxplots for responses to the three interest variables of SQ1, revealing a high interest in data sharing for all three variables.

³ Calculated using https://www.psychometrica.de/effect_size.html.

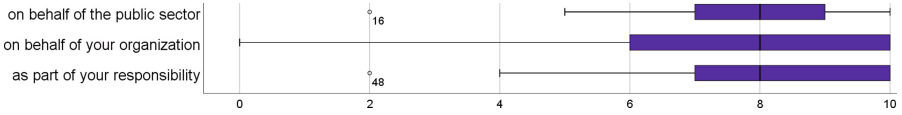


Fig. 1. IT practitioners' interest in data sharing (n = 61)

Pairwise tests for organizational levels indicate that divisional management is significantly more interested in data sharing *as part of their own responsibility* than are project/team leaders ($p = .035$, omnibus $d = .368$) and also significantly more interested in data sharing *on behalf of the public sector* than are specialists and experts ($p = .032$, omnibus $d = .511$).

Figure 2 shows boxplots for the three familiarity variables of SQ2, showing that familiarity with the possibilities and challenges of data sharing is closer to medium. Pairwise comparisons indicate that respondents feel they can contribute significantly less to *decisions regarding data sharing* than *explain data sharing in their own organization* ($p = .016$, omnibus $W = .100$).

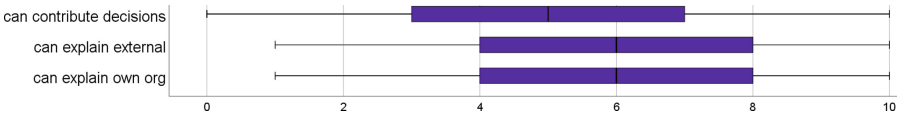


Fig. 2. IT practitioners' familiarity with possibilities and challenges with data sharing (n = 60)

The data exhibits significant and large differences across organizational levels for each of the three familiarity variables in Fig. 2 ($.006 \leq p \leq .023$, $.803 \leq d \leq .978$). Pairwise tests show that divisional managers tend to rate themselves as significantly better at explaining and making decisions about data sharing than do project and team leaders, specialists/experts, and to some extent, top managers ($.001 \leq p \leq .037$).

RQ2: The Contribution of Data Sharing to Selected Benefit Areas. Figure 3 gives boxplots for the eight benefits area variables of SQ3, showing that respondents perceive

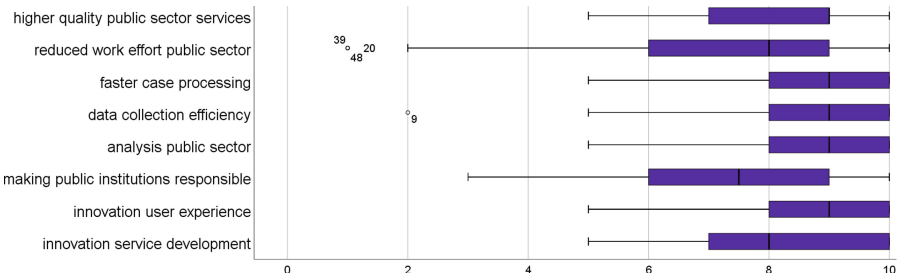


Fig. 3. Contribution of data sharing on benefits areas (n = 58)

the potential benefits from data sharing to be high or close to high for all benefits areas. The omnibus test across all eight variables reveals significant differences ($p = .000$) but with a small effect size ($W = .164$). Pairwise tests show that data sharing is perceived to benefit *making public institutions responsible* and *reduced work effort public sector* significantly less than all other benefits areas ($.000 \leq p \leq .011$). Similarly, data sharing is perceived to benefit *reduced work effort in the public sector* and *making public institutions responsible* significantly less than all other benefits areas ($.000 \leq p \leq .014$). Finally, data sharing is perceived to benefit *higher quality public sector services* significantly less than *improved analysis in the public sector* ($p = .020$).

Across respondents' organizational level, pairwise tests show that top management has a significantly higher ($p = .038$, omnibus $d = .385$) belief in a *reduction in work effort in the public sector* resulting from data sharing than do project or team leaders.

The omnibus test across all work domains shows a significantly large difference ($p = .038$, $d = .704$) in perceptions about *data collection efficiency*. Pairwise tests for work domains show that those working in technology have significantly higher expectations of *data collection efficiency* than do those working in development ($p = .013$) and those working in the line organization ($p = .041$).

RQ3: Value Creation in Public-Sector Segments. Figure 4 shows boxplots for the 15 public sector-segment variables of SQ4, where perceived potentials for value creation from data sharing are high to medium-high for all the segments. The omnibus test across all 15 variables shows a significant, small difference ($p = .000$, $W = .299$).

Pairwise comparisons show that *arts and culture* as well as *agriculture* are perceived to hold a significantly lower potential for value creation from data sharing than all the other variables ($.000 \leq p \leq .034$). Also, *research* is perceived to hold a significantly higher potential for value creation than all the other variables except for *across sectors*, *police and customs*, and *health* ($.000 \leq p \leq .038$), while *health* holds a higher potential than all except *welfare*, *police and customs*, and *across sectors* ($.000 \leq p \leq .040$). Other variables are also found to differ significantly, but against fewer variables.

RQ4: Impact of Challenges to Data Sharing: Figure 5 shows boxplots for the nine challenges variables of SQ5 which are perceived to have between medium and high impact. The omnibus test across all nine variables shows significant, small differences ($p = .000$, $W = .093$). Pairwise comparisons show that *a lack of top management support*, and to some degree *lacking goals/strategies*, and *technical competence* are considered less impactful than the other variables ($.000 \leq p \leq .027$). *Unfit technical infrastructure* is reported to have significantly less impact than *lacking common understanding and standards for data* ($p = .009$) and *lacking collaboration between organizations* ($p = .026$). *Lacking technical standards for collaboration* is reported to have significantly less impact than *lacking collaboration between organizations* ($p = .049$).

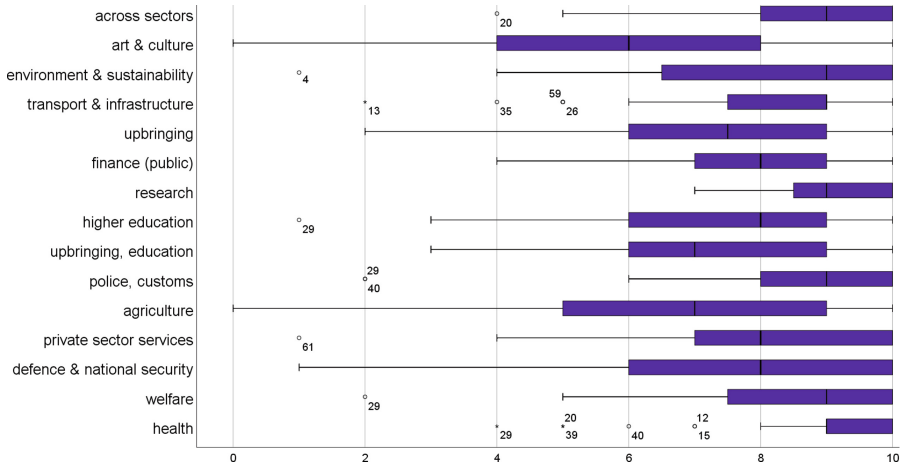


Fig. 4. Potential for value creation within public-sector segments (n = 56)

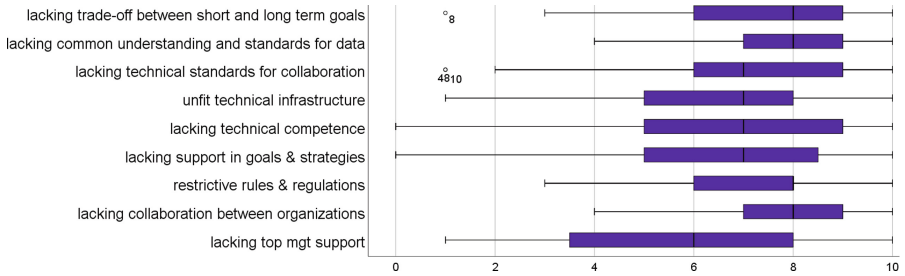


Fig. 5. Impact of challenges to data sharing (n = 55)

Pairwise comparisons on organizational level show that divisional management has significantly lower concern about *restrictive rules and regulations* than do project and team leaders ($p = .039$, omnibus $d = .355$). Divisional managers also have a significantly lower concern about *lacking trade-offs between short and long-term goals* than do specialists/experts ($p = .045$, omnibus $d = .504$).

Omnibus tests across work domains show significantly large differences in concerns about *lacking technical competence* ($p = .032$, $d = .734$) and *unfit technical infrastructure* ($p = .007$, $d = .949$). Pairwise comparisons indicate that there are different perceptions about the *impact of restrictive rules and regulations* (considered significantly lower by staff functions than development ($p = .025$), *lacking technical competence* (considered significantly lower by staff functions than technology ($p = .004$), and *unfit technical infrastructure* (considered significantly lower by staff functions than technology ($p = .025$) and development ($p = .001$)).

RQ5: Likely Funding Mechanisms for Data Sharing. Figure 6 shows boxplots for the three financing option variables of SQ7. Visual inspection shows that most funding

mechanisms are considered medium or above likely, with earmarked allocation being most likely, but with statistically insignificant differences.

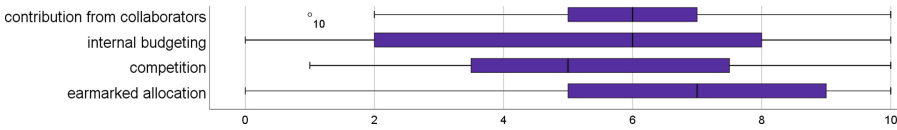


Fig. 6. Mechanisms for funding data-sharing solutions ($n = 56$)

RQ6: Confidence in the Public Sector to Realize Benefits and Overcome Obstacles. Figure 7 shows boxplots for the six requirements variables of SQ6. Visual inspection shows that practitioners’ faith in the public sector meeting requirements for data sharing is mostly around medium. The omnibus comparison across all the variables shows significant, small differences ($p = .012$, $W = .053$). Pairwise comparisons indicate that IT practitioners have lower faith in *learning from others’ experiences abroad* than *domestically* ($p = .027$), *understanding of impediments* ($p = .009$) and *understanding of benefits* ($p = .007$).

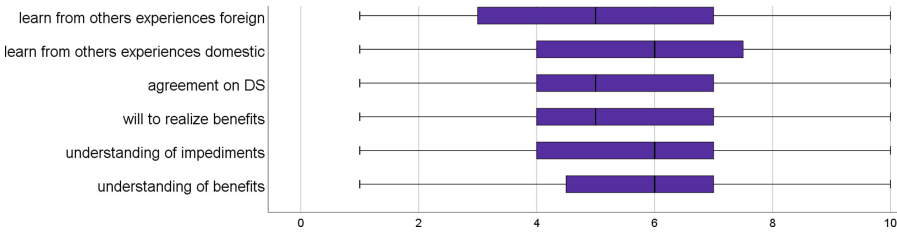


Fig. 7. Faith in the public sector meeting requirements for data sharing ($n = 55$)

Pairwise comparisons across respondents’ organizational level show that specialists/experts rate the public sector’s *understanding of impediments* as significantly lower than what divisional managers do ($p = .025$, omnibus $d = .473$). Top managers rate the public sector’s *will to realize benefits* as significantly lower than do specialists and experts ($p = .022$, omnibus $d = .519$).

Figure 8 shows boxplots for the two action variables of SQ8 and shows that the respondents’ perception of their own organization’s ability to realize the benefits of data sharing is medium, and the ability to handle impediments to data sharing is just above medium. No significant differences were found.

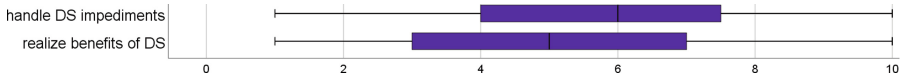


Fig. 8. Own organization’s ability in realizing benefits of, and handling impediments to, data sharing ($n = 56$)

6 Discussion

Respondents generally perceived significant benefits from sharing data, which was consistent with the optimism in the literature. However, the middling responses about concerns suggest uncertainty or ambivalence. Combining these with the low levels of confidence in the public sector’s ability to realize benefits and overcome obstacles indicates that data sharing solutions are still in early stages with a limited experience base. We do not yet have the basis to speculate why two types of benefits (public sector accountability and cost efficiency) and two segments (agriculture and arts/culture) were rated less promising for data sharing than the others, but is somewhat understandable in the light of ongoing public debate that health and research are rated highly as segments in which data sharing will have a positive impact.

Our data suggests that divisional managers see their responsibility differently than others do: they are more interested than others in data sharing, more confident in their understanding, and less concerned about obstacles than respondents at other organizational levels. Divisional managers may view data sharing as part of their responsibility. We expect this landscape to evolve in the next few years, most likely as part of a broader drive to integrate digitalization across public institutions.

7 Conclusion

Our findings about perceptions of the benefits of data sharing are consistent with the view that sharing data is an essential part of data-driven value creation. The optimism is tempered by misgivings about realizing the benefits and the lack of ability among public institutions to realize data sharing solutions.

In a broader sense, data sharing is a necessary component of a “dynamic system of systems” that enables innovative digitalization across organizations [1] – building awareness and capabilities about data sharing may be associated with the design and implementation of solutions that integrate across organizations.

8 Limitations

We provide the relevant information to replicate the survey so that other researchers/professionals can conduct it in other contexts. In the following, we present potential limitations for this study’s validity [18,37] of the study’s results and findings.

Construct Validity: For this exploratory survey, we developed concepts and categories by synthesizing themes from the literature to be used at the conceptual level in the

research questions. The questionnaire items were then designed with the intent to reflect those concepts. As described in Sect. 4, we evaluated the categories to avoid conceptual gaps and overlaps, also by getting feedback from external reviewers. Clearly, however, one should work further toward grounding the conceptual models empirically.

Internal Validity: By differentiating on grouping variables we believed to be relevant (i.e., respondents' level in the organization, their sector of employment, and work area), as well as their interest in and awareness of data sharing, we mitigated the threat of unstudied factors somewhat. Further comparative studies are needed when more is understood about what salient grouping factors may explain variations.

External Validity: An obvious threat is that the respondents are limited to the group of Norwegian IT practitioners present at the seminar. While their responses likely represent their roles in Norwegian public sector digitalization, we cannot be certain that their view applies to other roles and situations in other countries. We start with this small target audience to validate the suitability of the survey before conducting it in a broader context. We plan to conduct the survey at an international level to extend our dataset and substantiate our findings and comparisons further.

9 Implications for Research and Practice

Both our review of available literature and this survey suggest that data sharing is an emerging and important phenomenon that warrants further research. Hopes about benefits combined with concerns about obstacles, and particularly legal constraints, highlight both potential value and pitfalls for practitioners.

To this end, we hope that this paper provides the initial context and baseline for further research into data sharing, both in its own right and as part of the impetus for the public sector to become more data driven. Further, we suspect that the ability to build data-sharing solutions may reflect organizations' capability to digitalize across traditional divisions for the public good.

We also hope that this paper provides practitioners with better means to navigate issues related to data management, especially potential benefits and likely obstacles. Since the notion inherently calls for collaboration across public institutions, we believe that our findings may help facilitate productive discussions based on shared models and terminologies and that the work ahead to build solutions will enhance maturity in the field and accelerate learning.

References

1. Adler, R., et al.: A research roadmap for trustworthy dynamic systems of systems-motivation challenges and research directions. Technical report. IESE-001.23/E, Fraunhofer IESE, Germany (2023)
2. Aggarwal, A.K.: Opportunities and challenges of big data in public sector. In: *Web Services: Concepts, Methodologies, Tools, and Applications*, pp. 1749–1761 (2019)
3. Al-Khouri, A.M.: Strategy and execution: lessons learned from the public sector. *Int. Bus. Res.* 7(10), 61–73 (2014)

4. Bakken, I.J., Ariansen, A.M.S., Knudsen, G.P., Johansen, K.I., Vollset, S.E.: The Norwegian patient registry and the Norwegian registry for primary health care: research potential of two nationwide health-care registries. *Scand. J. Public Health* **48**(1), 49–55 (2020)
5. Basu, S.: E-government and developing countries: an overview. *Int. Rev. Law Comput. Technol.* **18**(1), 109–132 (2004)
6. Benfeldt, O., Persson, J.S., Madsen, S.: Data governance as a collective action problem. *Inf. Syst. Front.* **22**(2), 299–313 (2020)
7. Bigdeli, A.Z., Kamal, M.M., De Cesare, S.: Electronic information sharing in local government authorities: factors influencing the decision-making process. *Int. J. Inf. Manage.* **33**(5), 816–830 (2013)
8. Borgås, F.: General government revenue and expenditure. Technical report, Statistics Norway (2023)
9. Broomfield, H., Reutter, L.: In search of the citizen in the datafication of public administration. *Big Data Soc.* **9**(1), 205395172210893 (2022)
10. Broomfield, H., Reutter, L.M.: Towards a data-driven public administration: an empirical analysis of nascent phase implementation. *Scand. J. Public Admin.* **25**(2), 73–97 (2021)
11. Christodoulou, P., et al.: Data makes the public sector go round. In: Parycek, P., et al. (eds.) *EGOV 2018. LNCS*, vol. 11020, pp. 221–232. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98690-6_19
12. Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Routledge, Abingdon (1988)
13. Dawes, S.S.: Interagency information sharing: expected benefits, manageable risks. *J. Policy Anal. Manage.* **15**(3), 377–394 (1996)
14. Dener, C., Nii-Aponsah, H., Ghunney, L.E., Johns, K.D.: *GovTech Maturity Index: The State of Public Sector Digital Transformation*. World Bank Publications (2021)
15. Ebbing, M.: Et nytt system for enklere og sikrere tilgang til helsedata - rapport fra helsedatautvalget 2016–2017. Technical report, Helsedatautvalget til Regjeringen (2017)
16. Elvestad, L.A.: Deling av helseopplysninger gir økt pasientsikkerhet. *Dagens Medisin* (2019)
17. Favaretto, M., De Clercq, E., Schneble, C.O., Elger, B.S.: What is your definition of big data? Researchers' understanding of the phenomenon of the decade. *PLoS ONE* **15**(2), e0228987 (2020)
18. Feldt, R., Magazinius, A.: Validity threats in empirical software engineering research-an initial survey. In: Seke, pp. 374–379 (2010)
19. Gil-García, J.R., Helbig, N., Ojo, A.: Being smart: emerging technologies and innovation in the public sector. *Gov. Inf. Q.* **31**, I1–I8 (2014)
20. Gorry, G.A.: Sharing knowledge in the public sector: two case studies. *Knowl. Manage. Res. Pract.* **6**(2), 105–111 (2008)
21. Graham, F.S., Gooden, S.T., Martin, K.J.: Navigating the transparency-privacy paradox in public sector data sharing. *Am. Rev. Public Admin.* **46**(5), 569–591 (2016)
22. Guenduez, A.A., Mettler, T., Schedler, K.: Technological frames in public administration: what do public managers think of big data? *Gov. Inf. Q.* **37**(1), 101406 (2020)
23. Harvey, F., Tulloch, D.: Local-government data sharing: evaluating the foundations of spatial data infrastructures. *Int. J. Geogr. Inf. Sci.* **20**(7), 743–768 (2006)
24. Hautamäki, A., Oksanen, K.: Digital platforms for restructuring the public sector. *Collaborative value co-creation in the platform economy*, pp. 91–108 (2018)
25. Janssen, M., Charalabidis, Y., Zuiderwijk, A.: Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manag.* **29**(4), 258–268 (2012)
26. Janssen, M., Konopnicki, D., Snowdon, J.L., Ojo, A.: Driving public sector innovation using big and open linked data (BOLD). *Inf. Syst. Front.* **19**, 189–195 (2017)
27. Jean-Quartier, C., Rey Mazón, M., Lovrić, M., Stryeck, S.: Collaborative data use between private and public stakeholders-a regional case study. *Data* **7**(2), 20 (2022)

28. Jetzek, T., Avital, M., Bjorn-Andersen, N.: The sustainable value of open government data. *J. Assoc. Inf. Syst.* **20**(6), 6 (2019)
29. Karlsson, F., Frostenson, M., Prenkert, F., Kolkowska, E., Helin, S.: Inter-organisational information sharing in the public sector: a longitudinal case study on the reshaping of success factors. *Gov. Inf. Q.* **34**(4), 567–577 (2017)
30. McKight, P.E., Najab, J.: Kruskal-Wallis test. In: Weiner, I.B., Craighead, W.E. (eds.) *The Corsini Encyclopedia of Psychology*, vol. 1, pp. 1–10. Wiley, Hoboken (2010)
31. OECD: *Classification of the Functions of Government (COFOG)*. OECD (2021)
32. van Ooijen, C., Ubaldi, B., Welby, B.: A data-driven public sector: enabling the strategic use of data for productive, inclusive and trustworthy governance. *OECD Working Papers on Public Governance* (2019)
33. Pardo, T.A., Nam, T., Burke, G.B.: E-government interoperability: interaction of policy, management, and technology dimensions. *Soc. Sci. Comput. Rev.* **30**(1), 7–23 (2012)
34. Pardo, T.A., Tayi, G.K.: Interorganizational information integration: a key enabler for digital government (2007)
35. Pencheva, I., Esteve, M., Mikhaylov, S.J.: Big data and AI-a transformational shift for government: so, what next for research? *Public Policy Adm.* **35**(1), 24–44 (2020)
36. Ramon Gil-Garcia, J., Chengalur-Smith, I., Duchessi, P.: Collaborative e-government: impediments and benefits of information-sharing projects in the public sector. *Eur. J. Inf. Syst.* **16**(2), 121–133 (2007)
37. Runeson, P., Host, M., Rainer, A., Regnell, B.: *Case Study Research in Software Engineering: Guidelines and Examples*. Wiley, Hoboken (2012)
38. Saputro, R., Pappel, I., Vainsalu, H., Lips, S., Draheim, D.: Prerequisites for the adoption of the x-road interoperability and data exchange framework: a comparative study. In: 2020 Seventh International Conference on eDemocracy & eGovernment (ICEDEG), pp. 216–222. IEEE (2020)
39. Sawilowsky, S.S.: New effect size rules of thumb. *J. Modern Appl. Stat. Meth.* **8**(2), 597–599 (2009)
40. Tanilkan, S.S., Hannay, J.E.: Perceived challenges in benefits management-a study of public sector information systems engineering projects. In: 2022 IEEE 24th Conference on Business Informatics (CBI), vol. 1, pp. 156–165. IEEE (2022)
41. Tanilkan, S.S., Hannay, J.E.: Benefit considerations in project decisions. In: Taibi, D., Kuhrmann, M., Mikkonen, T., Klünder, J., Abrahamsson, P. (eds.) *Product-Focused Software Process Improvement*. LNCS, vol. 13709, pp. 217–234. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-21388-5_15
42. Thomas, M.A., Cipolla, J., Lambert, B., Carter, L.: Data management maturity assessment of public sector agencies. *Gov. Inf. Q.* **36**(4), 101401 (2019)
43. Tomar, L., Guicheney, W., Kyarisiima, H., Zimani, T.: Big data in the public sector: selected applications and lessons learned. Inter-American Development Bank, IADB (2016)
44. Tomczak, M., Tomczak, E.: The need to report effect size estimates revisited. an overview of some recommended measures of effect size. *Trends Sports Sci.* **1**(21), 19–25 (2014)
45. Turoń, K.: From the classic business model to open innovation and data sharing-the concept of an open car-sharing business model. *J. Open Innov. Technol. Market Complex.* **8**(1), 36 (2022)
46. Van Loenen, B., Grothe, M.: INSPIRE empowers re-use of public sector information. *Int. J. Spat. Data Infrastruct. Res.* **9**, 96–106 (2014)
47. van Veenstra, A.F., Timan, T.: A public value impact assessment framework for digital governance. In: Charalabidis, Y., Flak, L.S., Viale Pereira, G. (eds.) *Scientific Foundations of Digital Governance and Transformation*. PAIT, vol. 38, pp. 189–215. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-92945-9_8

48. Wilson, C., Mergel, I.: Overcoming barriers to digital government: mapping the strategies of digital champions. *Gov. Inf. Q.* **39**(2), 101681 (2022)
49. Wimmer, M.A., Neuroni, A.C., Frecè, J.T.: Approaches to good data governance in support of public sector transformation through once-only. In: Viale Pereira, G., et al. (eds.) *EGOV 2020. LNCS*, vol. 12219, pp. 210–222. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-57599-1_16
50. Yang, T.M., Maxwell, T.A.: Information-sharing in public organizations: a literature review of interpersonal, intra-organizational and inter-organizational success factors. *Gov. Inf. Q.* **28**(2), 164–175 (2011)
51. Zhang, J., Dawes, S.S., Sarkis, J.: Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing. *J. Enterp. Inf. Manage.* **18**(5), 548–567 (2005)


Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Investigating the Barriers that Women Face in Software Development Teams Focusing on the Context of Proprietary Software Ecosystems

Juliana Carvalho Silva do Outão¹(✉) , Luiz Alexandre Martins da Costa¹, Rodrigo Pereira dos Santos¹, and Alexander Serebrenik²

¹ Universidade Federal do Estado do Rio de Janeiro, Av. Pasteur, 458, Rio de Janeiro, Brazil

juliana.carvalho@edu.unirio.br

² Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

Abstract. Despite the growing discussion and concern about the topic, gender diversity in the Exact Sciences and Technology still requires attention. It has been observed by several authors that gender diversity is not present in a significant way in development teams, despite the potential positive effects. Moreover, with the growing demand for software that meet complex business needs, the concept of Software Ecosystems (SECO) has emerged and opens opportunities for external developers and strategies for fostering gender diversity. A Proprietary Software Ecosystem (PSECO) is a type of SECO that comprises a common technological platform with contributions protected by intellectual property. This work aims to investigate which barriers women face in software development teams focusing on the context of PSECO and what strategies can be used to increase inclusion based on a multivocal literature review. To do so, 29 studies were selected and 13 gender barriers were identified, with the 3 most cited barriers being: sexism, lack of peer parity, and imposter syndrome. Furthermore, it was observed that external PSECO actors can significantly interfere in the occurrences of gender barriers, in addition to the internal actors of the central organization (keystone).

Keywords: Diversity · Human Factors · Proprietary Software Ecosystems

1 Introduction

A significant gender disparity, with women being underrepresented, can be observed in the software industry [7]. Research has also shown that gender diversity in corporate boardrooms positively influences market value and profitability [1]. This underrepresentation of women in the software industry and development teams is attributed to persistent barriers that hinder diversity.

The Information and Communication Technology sector has been growing at a fast pace in recent years [3]. This sector traditionally demands a large number of

professionals in the areas of Science, Technology, Engineering and Mathematics (STEM) who are mostly male professionals. In recent years, the development of new, modern, and innovative systems that meet the ever-expanding business needs has become a challenging task for companies. From this need, software ecosystems (SECO) emerge as a solution to deal with such scenario [2]. The type of SECO in which the value creation is based on proprietary contributions, protected by intellectual property management processes, is called Proprietary SECO (PSECO). In PSECO, where actors and their relationships are key roles, investigating gender diversity is also important for the environment.

In this context, the present study aims to identify the barriers that women face in software development teams in a PSECO context. Thus, a Multivocal Literature Review (MLR) was conducted to identify gender barriers and strategies to deal with such barriers, from the point of view of academia and industry.

2 Research Method

MLR emerged in the early 1990s, combining Systematic Literature Reviews (SLR) and Systematic Mapping Studies (SMS) that encompass both academic and gray literature [9]. This approach was chosen because many software professionals do not publish in academic forums, making the inclusion of gray literature essential to capture their insights. Gender diversity is a prominent industry topic, offering valuable perspectives. We followed the MLR model by Garousi et al. [6], which is rooted in Kitchenham and Charters' guidelines for SLR and SMS [8]. Protocol development and application took place between November 2022 and September 2023.

To address the purpose of the study, the following main research question (RQ) was defined: **What are the barriers to gender diversity in software development teams and what are the strategies to deal with such barriers focusing on the proprietary software ecosystem context?** To answer the RQ, the following sub-questions (SQ) were elaborated: **(SQ1) What are the barriers that women face in software development teams?**; and **(SQ2) What are the strategies to foster gender diversity in software development teams?**. After some refinements, the following search string below was used and Fig. 1 illustrates an overview of the process: **(women OR “gender diversity” OR “gender inclusion” OR “gender equity” OR “gender equality” OR “gender bias”) AND (“software engineering” OR “software ecosystem” OR “software development” OR “open source” OR “software industry”) AND (barrier* OR challenge* OR issue*)**

Unlike the scientific literature, determining when to conclude an MLR is complex due to the number of substantial results. In this study, we adopted the limited effort criterion based on Garousi et al.'s guidelines [6]. We assessed the first 100 search results for each database (200 studies in total), continuing the search only if the last page showed potential relevant findings. After examining the next page following the initial 100 records, no additional studies were deemed suitable for inclusion in the MLR.

3 Results

After executing the MLR process described in Sect. 2, information was extracted from 29 selected studies, which were numbered from S01 to S29. Further details about the selected studies are available via Zenodo¹. To respond the main RQ, the both SQ were answered, as described next. It is noteworthy that encodings were performed based on the qualitative analysis from Grounded Theory procedures [10].

SQ1 - What are the Barriers that Women Face in Software Development Teams? Applying code procedures, 13 gender barriers were identified from the selected studies. Details on the identified barriers and the number of studies for each barrier is described bellow. It is noteworthy that a study may have described one or several barriers. To assist in their understanding, the definition of each barrier is described below:

1. **Sexism (identified in 16 studies):** Sexism can be hostile or benevolent. Hostile sexism is prejudice itself (microaggressions), such as not being heard in technical discussions and receiving derogatory comments that women perform inferiorly to men. In turn, benevolent sexism represents subjectively positive feelings towards a gender that often brings some sexist antipathy, reinforcing the idea that women need to be cared for by men;
2. **Lack of peer parity (identified in 15 studies):** Peer parity is the concept that an individual can identify herself/himself with at least one other peer when interacting in a community;
3. **Imposter syndrome (identified in 14 studies):** Individuals who experience intense feelings that their achievements are undeserved and fear that they may be exposed as frauds;
4. **Technical difficulties (identified in 10 studies):** This barrier refers to technical problems, such as lack of knowledge, lack of experience, and unfamiliarity with the technology or programming language used;

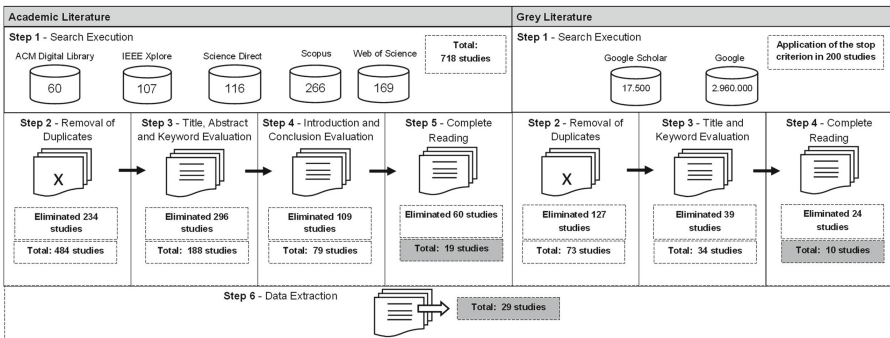


Fig. 1. Process applied in MLR.

¹ <https://doi.org/10.5281/zenodo.10056419>.

5. **Non-inclusive communication (identified in 8 studies):** This barrier refers to the use of exclusionary communication, such as the use of profanity and terms generally associated with men (for example, “guys”);
6. **Imbalance between personal and professional life (identified in 8 studies):** This barrier refers to the lack of support for well-being, causing an imbalance in the personal lives of professionals. It is usually caused by too much overtime or pressure to deliver activities;
7. **Stereotypes (identified in 8 studies):** Stereotypes are beliefs about characteristics, attributes and behaviors of certain members of a group;
8. **Prove it again (identified in 7 studies):** It refer to the bias effect that occurs when a member of a group does not align with the stereotypes is measured to a higher standard and has to provide more evidence to demonstrate competence;
9. **Harassment (identified in 6 studies):** Harassment is abusive conduct demonstrated by means of words, behaviors, acts, gestures, or writings that may harm a person’s personality, dignity or physical, or mental integrity, endangering their employment, or degrade the work environment;
10. **Glass ceiling (identified in 5 studies):** This is a transparent barrier that prevents women from rising above a certain level in corporations;
11. **Lack of recognition (identified in 4 studies):** Not feeling valued and not being recognized when good work is done;
12. **Toxic culture (identified in 4 studies):** It is characterized by work environments where there is room for favoritism, rumors, and people trying to harm each other;
13. **Maternal and family issues (identified in 4 studies):** Describes the experience of women who have children or someone in their family who requires care and suffer prejudice due to this situation, being excluded from certain opportunities.

SQ2 - What are the Strategies to Foster Gender Diversity in Software Development Teams?

Based on the selected studies, it was possible to identify some strategies to foster gender diversity in software development teams. Most of the items listed below were identified in S13, which brought a detailed analysis of how to address each of the challenges mapped in its study. Below is a breakdown of the 7 identified high-level strategies and 26 actions to address each of them:

1. Embrace equality: give training to all managers regarding soft skills to be more empathetic and avoid burnout (Ac.01); respect and give voice to women (Ac.02); ensure equal pay (Ac.03); provide opportunities and challenges (Ac.04); not allocate women only to operational tasks (Ac.05); and give career choices to women in the same rate as men (Ac.06);
2. Supporting women’s career growth: encouraging women to advance in their careers (Ac.07); have more women in (technical) leadership (Ac.08); and mentor other women who are role models (Ac.09);
3. Support work-life balance: implement well-being policies (Ac.10); discourage overtime (Ac.11); improve location and time flexibility (Ac.12); and support parenthood (Ac.13);

4. Empower Women: publicize women's successes on social media and at external events (Ac.14); and recognize and reward women's achievements (Ac.15);
5. Hire more women: make job opportunities attractive to women's needs (Ac.16); change the recruitment and marketing processes (Ac.17); have more women recruiting for open positions (Ac.18); create IT vacancies aimed exclusively at women (Ac.19); and invest in programs to attract girls to STEM (Ac.20);
6. Promote women's groups and events: organize supporting groups for women (Ac.21); promote interaction between women (Ac.22); and organize campaigns and/or lectures on the importance of gender diversity (Ac.23);
7. Create and reiterate policies: create, disseminate, and raise awareness of the code of conduct (Ac.24); promote anti-harassment policies (Ac.25); and make explicit statements that there is zero tolerance for anti-gender inclusive behavior (Ac.26).

4 Discussion

In the bibliometric analysis, recent studies, primarily from the United States, were selected, with 2022 and 2018 having the most publications. Notably, the most frequently cited barriers in the selected studies were sexism, lack of peer parity, and imposter syndrome. Trinkenreich et al.'s study [11] on women in open-source software communities also highlighted imposter syndrome and lack of peer parity as key barriers. This study additionally identified seven other barriers, including harassment, technical difficulties, glass ceiling, lack of recognition, and maternal and family issues.

Analyzing these results in the context of PSECO, the barriers were categorized into internal and external barriers. Internal barriers included imposter syndrome and maternal and family issues, while external barriers encompassed sexism, lack of peer parity, glass ceiling, lack of recognition, non-inclusive communication, prove it again, imbalance between personal and professional life, technical difficulties, stereotypes, harassment, and toxic culture.

Despite PSECO having its own characteristics, developers interact with other actors through ecosystem relationships. External barriers apply to PSECO, addressing interactions with keystones or other ecosystem actors. However, internal barriers should not be overlooked and require proper evaluation for inclusive environments.

An SLR performed by Canedo et al. [5] highlighted strategies to increase women's participation in open source projects, similar to those found in the present study, such as exclusive vacancies for women, training, code of conduct, and inclusive policies. Continuous monitoring of female participation for metrics generation was also suggested. Van Breukelen [4] emphasized the intersection between multiple minority groups, such as veteran women or black women, who face unique barriers, requiring targeted strategies for meaningful change.

5 Final Remarks

We conducted an MLR to explore gender barriers in software development teams within the PSECO context, revealing 13 gender barriers in total, but 11 distinct barriers that are beyond the organizational boundaries, involving external actors such as clients and suppliers. We also identified strategies to address these gender barriers and promote women's inclusion in this environment.

Regarding threats to validity, our study covered specific databases and some grey literature was not evaluated, but we followed recommended stopping criteria. We acknowledge that our search was limited to English-language studies, but this aligns with the prevalent language in global academic research.

To mitigate potential bias, we discussed inclusion criteria with other researchers and conducted a thorough review process. In future work, a field study could validate the identified barriers among women in real PSECO settings. Additionally, similar MLR studies could be conducted to map barriers and strategies for other types of diversity beyond gender with a focus on women.

References

1. Arnoğlu, E.: Female board members: the effect of director affiliation. *Gend. Manage.* **35**, 225–254 (2020)
2. Barbosa, O., Santos, R., Alves, C., Werner, C., Jansen, S.: *A Systematic Mapping Study On Software Ecosystems from a Three-dimensional Perspective*. Edward Elgar Publishing, Cheltenham (2013)
3. Botella, C., Rueda, S., López-Iñesta, E., Marzal, P.: Gender diversity in stem disciplines: a multiple factor problem. *Entropy* **21**(1), 30 (2019)
4. van Breukelen, S., Barcomb, A., Baltés, S., Serebrenik, A.: “still around”: experiences and survival strategies of veteran women software developers. In: *Proceedings of the 45th International Conference on Software Engineering, ICSE 2023*, pp. 1148–1160. IEEE Press (2023)
5. Canedo, E., Tives, H., Bogo, M., Fagundes, F., Cerqueira, J.: Barriers faced by women in software development projects. *Information* **10**, 309 (2019)
6. Garousi, V., Felderer, M., Mäntylä, M.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Inf. Softw. Technol.* **106**, 101–121 (2019)
7. Hill, C., Corbett, C., Rose, A.: *Why so few? Women in science, technology, engineering, and mathematics*. American Association of University Women, January 2010
8. Kitchenham, B., Charters, S.: *Guidelines for performing systematic literature reviews in software engineering*. Technical report. EBSE 2007–001, Keele University and Durham University Joint Report (2007)
9. Ogawa, R., Malen, B.: Towards rigor in reviews of multivocal literatures: applying the exploratory case study method. *Rev. Educ. Res.* **61**, 265–286 (1991)
10. Strauss, A., Corbin, J.: *Grounded Theory: Methodology Applied in Education Research*. Artmed (2008)
11. Trinkenreich, B., Wiese, I., Sarma, A., Gerosa, M., Steinmacher, I.: Women's participation in open source software: a survey of the literature. *ACM Trans. Softw. Eng. Methodol.* **31**(4), 1–37 (2022)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.












The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Artificial Intelligence



Business and Ethical Concerns in Domestic Conversational Generative AI-Empowered Multi-robot Systems

Rebekah Rousi¹  , Hooman Samani² , Niko Mäkitalo³ , Ville Vakkuri¹ ,
Simo Linkola⁴ , Kai-Kristian Kemell⁴ , Paulius Daubaris⁴ , Ilenia Fronza⁵ ,
Tommi Mikkonen³ , and Pekka Abrahamsson⁶ 

¹ University of Vaasa, Wolffintie 32, 65200 Vaasa, Finland
rebekah.rous@uwasa.fi

² University of Arts London, 272 High Holborn, London WC1V 7EY, UK

³ University of Jyväskylä, Mattilanniemi 2, 40100 Jyväskylä, Finland

⁴ University of Helsinki, Yliopistonkatu 4, 00100 Helsinki, Finland

⁵ Free University of Bozen-Bolzano, Sparkassenstraße 21 - via Cassa di Risparmio, 21 39100
Bozen-Bolzano, Italy

⁶ University of Tampere, Pohjoisranta 11A, 28100 Pori, Finland

Abstract. Business and technology are intricately connected through logic and design. They are equally sensitive to societal changes and may be devastated by scandal. Cooperative multi-robot systems (MRSs) are on the rise, allowing robots of different types and brands to work together in diverse contexts. Generative artificial intelligence has been a dominant topic in recent artificial intelligence (AI) discussions due to its capacity to mimic humans through the use of natural language and the production of media, including deep fakes. In this article, we focus specifically on the conversational aspects of generative AI, and hence use the term Conversational Generative artificial intelligence (CGI). Like MRSs, CGIs have enormous potential for revolutionizing processes across sectors and transforming the way humans conduct business. From a business perspective, cooperative MRSs alone, with potential conflicts of interest, privacy practices, and safety concerns, require ethical examination. MRSs empowered by CGIs demand multi-dimensional and sophisticated methods to uncover imminent ethical pitfalls. This study focuses on ethics in CGI-empowered MRSs while reporting the stages of developing the MORUL model.

Keywords: Multi-robot cooperation · Business · Ethics · Conversational Generative AI · Large Language Models

1 Introduction

Generative Artificial Intelligence is currently in the spotlight, drawing both praise and criticism. Conversational AI, on the other hand, has been studied for several years and refers to chatbot technologies which are somehow considered to make the interactions

with the chatbot intelligent. In this article, we use the term Conversational Generative Artificial Intelligence (CGI) to refer specifically to the combination of generative and conversational artificial intelligence (AI). It has permeated every corner of society, revolutionizing communication between humans and machines using natural language. Two fields significantly impacted by this technology are business and robotics. Integrating CGI into organizational operations can yield substantial business value [1]. Similarly, employing CGI in robotics enhances usability, accessibility, and the market potential of robotic systems [2]. However, embracing these cutting-edge technological developments is not without risks. Recent headlines in major media outlets have underscored the potential consequences of mishaps in sophisticated data-driven systems for humans, technology, and businesses alike.

One of the primary contexts for deploying these complex emerging products and services is the home. For instance, the global smart home market is projected to grow from \$93.98 billion in 2023 to \$338.28 billion by 2030 [3]. This rapid growth in the market introduces a complex landscape, integrating multi-layered Systems of Systems (SoSs) into the traditionally private and sacred space of the home [4, 5]. Everyday products such as refrigerators, vacuum cleaners, and toasters are transforming into intelligent devices with the potential to function as discreet communicators [6]. Consequently, ethical considerations are intertwined with all levels of technological implementation in the home due to the changing dynamics in human-object relationships [7].

The presence of CGI-embedded Multi-Robot Systems (MRS) in domestic settings raises a multitude of ethical concerns for businesses [8, 9]. The development of CGI-embedded MRSs has predominantly focused on industrial and business applications [10]. These systems aim to automate tasks and enhance efficiency in various industries, including manufacturing, healthcare, and customer service. As a result, the ethical dimensions of CGI-embedded MRSs have often been overlooked. Businesses engaged in the development or deployment of CGI-embedded MRSs must carefully consider these ethical concerns and take steps to address them. This paper adopts an applied ethics approach to explore potential ethical issues arising from the development and deployment of data-driven multi-robot cooperative systems. Applied ethics, in this context, refers to a case-specific approach that examines how social ethical dilemmas manifest practically when specific technical and social-technical elements (involving a blend of human and technological factors) are put into operation in specific contexts [11].

Instead of seeking to already *solve* problems, this study primarily focuses on *identifying* potential ethical challenges during the development, deployment, and implementation of multi-robot cooperative systems for implementation in the home. As this is a novel context in the area of AI ethics, we consider such problem identification important at this stage. In this respect, we consider the concept of moral awareness essential in order to go beyond the concerns voiced in existing literature on AI ethics. *Moral awareness* is defined as the ability to identify ethical aspects in a given context [12]. In this paper, a scenario-based approach is employed to investigate the potential ethical concerns and moral implications of introducing heterogeneous multi-robots into domestic spaces.

More specifically, the authors aim to develop a model for promoting moral awareness in multi-robot systems (MRSs) – the MORUL model. Furthermore, the authors recognize

that not all ethical issues and related interventions can be addressed during the pre-development phases. In the emerging MORUL model, ethical concerns are mapped and predicted in relation to stages at which analyses should be conducted. These analyses are carried out with regard to the dimension affected by the ethical concern, such as safety, security, or societal impact. This paper contributes to and builds upon previous efforts that sought to establish ethical practices and frameworks for the development of artificial intelligence (AI) [13].

2 Background

2.1 Large Language Models (LLMs) in Multi-robot Cooperation

Large Language Models (LLMs) and Generative Artificial Intelligence represent some of the latest developments in machine learning that have gained widespread public attention. OpenAI's Generative Pre-training Transformer architecture (ChatGPT) has been at the center of headlines and public debates since around 2018 [1]. LLMs are part of the recent trend in the growing popularity of chatbot development [14], which make Conversational Artificial Intelligence stand out as an advancement towards higher AI development goals such as Artificial General Intelligence (AGI). Hence, we use the term 'Conversational Generative Artificial Intelligence' (CGI) in this article to be specific about the technology we are referring to. In the case of chatbots, Natural Language Processing (NLP) is employed to interact with users by providing optimal responses from the information system. ChatGPT can be viewed as an advanced form of chatbot, enhancing earlier versions by combining deep learning and LLMs [15]. LLMs focus on predicting word sequences commonly used in human communication. However, this process introduces biases and discrimination due to the reliance on neural network transformer architectures and deep learning, which depend on representative data [16]. For instance, ChatGPT combines supervised fine-tuning with unsupervised pre-training to generate responses that appear to be human-like, thus expanding the social dimension of human-data interaction and improving data accessibility for non-experts.

Currently, engaging in prompt-based conversations with AI-based chatbots can be relatively expensive, considering the number of prompts typically required for a single task and the widespread usage of these models. Tech companies like OpenAI, Microsoft, Alphabet, and Meta are striving to capitalize on this emerging technology by building businesses around AI-based applications for personal and professional use. Given the costs associated with training and running these models, companies are competing with diverse business strategies. OpenAI, for example, offers its GPT model as a service via an API, allowing new AI-based applications to be developed on top of their models. Meanwhile, new open-source LLMs with various capabilities and licenses are being released on the internet. Meta, for instance, provides its advanced LLAMA 2 model as open source, with limited commercial use.

Multi-robot cooperation involves two or more robots, regardless of brand, model, or type, working together to achieve shared goals [17]. While each robot may have unique objectives, there should be a common overarching goal among them, such as ensuring a safe and clean home or delivering timely and effective services in a hospital. The ultimate goal in such scenarios is typically the well-being of the human owner.

Multi-robot cooperation primarily addresses complex tasks that are nearly impossible to accomplish successfully without a team effort [17, 18]. At all stages, human involvement is a constant factor, whether it's in programming, giving commands, or collaborating with the robots. Consequently, multi-robot cooperation should always be considered in relation to humans and their varying levels of involvement in different processes [19]. Considering human factors in working with multi-robot systems introduces different levels of complexity, as identified by Simões and colleagues [20]: 1) the human operator and the technology itself; 2) recommendations and guidelines affecting the performance of human-robot teams; and 3) complex holistic approaches guided by recommendations and guidelines that influence human-robot interaction.

In any case, it is essential to recognize that the human dimension in multi-robot cooperation is always the result of complex negotiations between integrated systems, diverse operational goals, varied corporate strategies, governed by standards, laws, and recommendations. Therefore, the starting point for examining such systems always begins at Level 3 [20]. Preempting ethical issues during the pre-development phase elevates the investigation to Level 4, involving systemic ethical forecasting in cybernetic systems. This forecasting requires an understanding of how Multi-Robot Systems (MRSs) operate within human contexts, with communication playing a crucial role [21]. Communication not only involves the functional aspects of human interaction with multi-robot systems but also encompasses the social-emotional components of Human-Robot Interaction (HRI) [21, 22]. As a result, CGI in forms such as ROSGPT or ChatGPT has significantly impacted the ways people interact with machine learning systems [23].

ROSGPT [24] introduces an innovative approach that leverages the full potential of LLMs to enhance human-robot interaction significantly. This framework integrates ChatGPT into ROS2-based robotic systems, creating a synergy between language understanding and robotic control. ROSGPT's advantage lies in its effective prompt engineering, utilizing ChatGPT's versatile capabilities, from information elicitation to coherent train of thought, to convert unstructured natural language commands into precise, contextually relevant robotic instructions. ROSGPT capitalizes on the inherent learning capabilities of LLMs to effortlessly extract structured commands from unrefined language inputs. The proof-of-concept demonstration, highlighting the translation of human language into actionable robotic instructions, underscores ROSGPT's potential across a range of applications. Beyond its immediate utility, ROSGPT's open-source implementation on ROS 2's platform not only fosters collaboration between the robotics and natural language processing fields but also represents a significant step toward the realm of AGI.

2.2 Business Effects of AI Ethics, CGI and Multi-robot Cooperation

Ethics in the domains of AI have been hot topics for decades now, and this is becoming increasingly more so as AI is deployed widely in society. Earlier discussions applied the terms 'information ethics', 'machine ethics' and 'computer ethics' [13, 25] to describe the field of examining ethical and moral implications of IT. With the broadening adoption of AI technologies in a multitude of domains, various practical incidents have highlighted diverse risks associated with AI.

The existing discussion on AI ethics, which far predates recent incidents, has served to identify and understand many of the risks already in the past - before they unfolded in actuality. Now, these predicted risks are becoming real, meaning that they present practical issues enabled by recent progress in ML. These risks are typically approached in research and development through *principles* in AI ethics [13]. For instance, racism, which is often associated with the principle of fairness, not only manifests through abuse and degradation, but also false accusation (see e.g., [26]). There is a sense of urgency spurred from the already emergent incidents involving machine learning (ML) technology utilization [25]. Whether the incidents involve matters of accountability and responsibility as witnessed in accidents in which human life has been harmed or damaged. The AI Incident Database [26] reported 90 incidents in 2022 alone, of AI-caused accidents, 45 already at the beginning of 2023. The rate of AI incidents seems to be increasing at a comparative pace to Moore's Law - doubling every year, similarly to the compounding capacity of computing speed [27]. These not only incur substantial costs in damages and potential insurance premiums, but pose serious problems from basic issues of human respect, safety, and dignity, to the severe tarnishing of reputation for businesses who do not embrace humane factors as a part of their data-driven business strategy [28].

The 2018 self-driving Uber accident in which a pedestrian was fatally wounded (see e.g., [29]) incurred irreparable immaterial damage. This no doubt contributed to loss of income, hindered self-driving vehicle development (and brands), tarnished Uber (now owned by Aurora Innovations) as a transportation service, and the operator who was responsible for monitoring the vehicle. While the human operator has been found guilty of negligence, the repercussions of the accident in terms of legal expenses and loss of consumer trust are remarkable. Not only were the direct implicated actors affected, but the US Federal Government was also accused of not properly regulating the industry. Moreover, had the accident led to a total abandonment of self-driving vehicles by companies such as Uber, profit trajectories would be thrown off course, because drivers account for 80% of all costs - self-driving units being evaluated at 7 billion United States dollars already in 2020 [3029].

Business intercedes on many dimensions of AI and robot ethics. From privacy-related issues and dark practices of the surveillance economy, to platform economy logic, and 'login - lock-in' cultures, business needs to be considered from both back and front-end perspectives. When it comes to ethics, business itself can be its own worst enemy. The logic that may pave the way to patents and trade secrets, may be guilty of fostering ethical potholes such as black box systems diminishing customer and user trust, and even simply, bad user experience with greater social repercussions. The dance between ethics and business is like a temptation-filled devil's tango. The appeal of fast profits blinds many of careful foresight in business strategy. Effective management of ethics in AI and robotic development would not just mean better business strategy, but also longevity [31].

3 Method

In the present study the researchers employed a qualitative exploratory method via two workshops. A scenario-based approach was used to contextualize the inquiry that entailed imagining that several robots of different use purpose, brand and type, utilizing CGI technology were implemented in the home (see Fig. 1). In the scenario, two cleaning robots of the same brand and make have been used in the home for quite some time. The new addition of a robot arm from a different brand and manufacturer elicits ethical concerns when considering the need for all robots to cooperate in order to perform tasks to reach certain goals. The goal of the workshops was to spark moral awareness in the participants in order to recognize ethical concerns and compare the identified concerns to those existing via previous research, and found in AI ethics guidelines and principles. The workshops were held at separate times: Workshop 1 (W1) was held during February, 2023, for two days face-to-face at a lab hosted by one of the participating research institutions; and Workshop 2 (W2) was held in June, 2023, for one hour via Zoom. The idea behind the separate timing was to allow for the analysis of W1 results, in order to synthesize and construct a preliminary framework for W2. The preliminary framework was seen as the basis for modeling a matrix that eventually will serve as a scaffolding for ethical multi-robot development. The matrix would include facets starting from ethical business strategy (understanding the influence of economic superstructures in molding the logic of technological products), to hardware and software, human-technology interaction, larger societal repercussions, and back again to business impact.

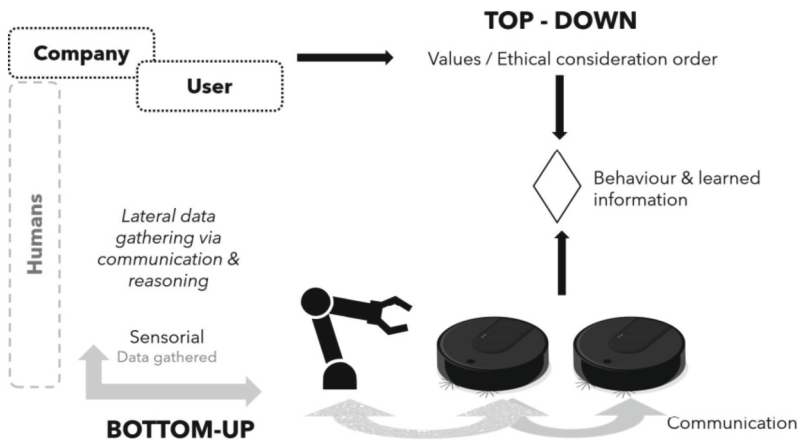


Fig. 1. Domestic scenario of two cleaning robots and one robot arm - understanding relations between layers and domains of multi-robot cooperation from a techno-corporate perspective

Qualitative data was collected in the form of brainstorming drawings and notes. The material from W1 was originally in paper versions, which were subsequently photographed and digitally archived. The material from W2 was produced on Google Jamboard. During processing of the data - transferal from the drawing boards to excel and image files - preliminary thematic categories were established. Extra rounds of thematic

analysis [32] were performed by the research team in an excel document. The study was conducted via a constructivist grounded theory [33] approach in order to build on previous AI ethics principles, guidelines and methods (see e.g., [18], while allowing for deeper examination of specific details and dimensions that are phenomenologically unique to the domain of multi-robot cooperation.

3.1 Ethical and Responsible Research

As this is a novel space of research that deals with ethics across a range of levels, from basic practical levels to higher levels of abstraction, the research team deemed the safest and most responsible approach to be that of internal inquiry. To avoid physical or psychological harm, the team of experts maintained the empirical component outside the realm of physical human-robot or robot-robot interactions. Rather, the researchers deliberated through discussion, illustration and writing. All researchers involved in the workshops were willing participants, agreeing the use of their data, exercising scholarly agency as experts within their respective fields. In compliance with the General Data Protection Regulation (GDPR), all data is stored in secure password-protected digital locations to which only two main researchers have access. No personal data is stored with the research data.

3.2 Participants

Each workshop comprised eight participants, rendering $N = 16$ contributions in total. Five participants participated in both workshops ($N = 10$ contributions) while six participants only participated in one of the workshops. This meant that the overall total of individual participants was $N = 11$. All participants possessed a higher tertiary degree, starting at PhD level researchers and higher. The gender distribution was two females and nine males. The fields of expertise that the participants represent are: software engineering and computer science; robotics and software for robotics; edge intelligence; computing education; information systems; cognitive science; human computer interaction; communication; and social ethics.

3.3 Procedure

The workshops were planned and agreed upon in a series of online meetings. In these meetings the strategy was deliberated, goals were set, as well as timing, procedure and locations were established. The context for the scenario was decided upon via several brainstorming sessions in which the team examined areas, environments and situations in which ethics and moral conduct would be considered as most sensitive [5]. After identifying several domains including education, healthcare, elderly care, and the home, the team selected the home, both for its intimate framing of privacy, as well as its diversity [4]. While there are central features defining a home - living space, kitchen, bedroom etc. - the ways in which people appropriate, populate, and utilize their spaces is quite

eclectic [4]. This is as opposed to public institutions such as hospitals that are laden with rules, standards and top-down regulations.

Workshop 1

Workshop 1 took place in person, on location at the lab of one of the participating research institutions. The lab is designed as an innovation space with a central meeting area equipped with audio-visual and teleconferencing equipment, as well as traditional tools such as flipcharts, post-it notes, colored pens. One participant contributed via Zoom for logistical reasons. The workshop was held over a two-day interval. The procedure entailed a round of introductions and articulating our interests in relation to the topic for the participants who had not been involved in the previous online planning sessions. The workshop proceeded as seen in Table 1.

Table 1. Workshop 1 procedure.

Step No.	Step label	Description
1	<i>Re-cap of use context and scenario</i>	Narrative unfolds in the home. Two similar robots (vacuum cleaners) and a newly introduced robot arm
2	<i>Independent mind-mapping of ethical concerns [unstructured]</i>	Independent work (30 min.), focus on ethical concerns
3	<i>Group discussion and comparison of findings</i>	Discussion of mind-maps, sharing ideas and introducing new concerns that arose in the group discussion
4	<i>Identification of the layers</i>	Identifying layers implicated in LLM-enabled multi-robots
5	<i>Model formulation</i>	Deliberation of actionable models of ethics in multi-robot collaboration that could be utilized within the programming process

Workshop 2

Workshop 2 was carried out via Zoom to allow for international collaboration while some members of the study were traveling. The duration of the workshop was two hours and held on Google Jamboard. Building on the findings of Workshop 1, Workshop 2 was structured according to a matrix of multi-robot cooperation domains and layers: Human-Interaction; Sensorial Layer (robot hardware); Deliberation (robot brain); Behavioral (robot hardware); Communication and Networking (robot-to-robot interaction); and System of Systems (network or systems). From the human perspective, considerations of ethical aspects were encouraged to be thought of through the frames of: 1) safety, 2) security, and 3) societal dimensions. The procedure of Workshop 2 is observed in Table 2.

Table 2. Workshop 2 procedure.

Step No	Step label	Description
1	<i>Instructions & breakdown of procedure + use-context recap</i>	Use context is the home and workshop members are encouraged to think of all potential ethical issues and scenarios arising from the introduction of LLM-powered multi-robot cooperation in domestic spaces
2	<i>Independent mind-mapping of ethical concerns [unstructured]</i>	Independent work (30 min.), focus on ethical concerns
3	<i>Group discussion & comparison of findings</i>	Groups progressed through the domains and layers of multi-robot cooperation as well as the human dimensions of the concerns
4	<i>Layer and domain refinement</i>	Group reflected on the earlier version of the layers and domains based on new findings arising in W2
5	<i>Model refinement</i>	MORUL model for ethical CGI-enabled multi-robot development further refined

3.4 Analysis

Thematic analysis [32] was employed to analyze the data of both workshops. In the case of Workshop 1, the researchers transcribed mind-maps, notes and illustrations that had been expressed on large flip chart sheets into excel sheets. From Workshop 2, the Google Jamboard notes were transferred into excel. The analysis took place in three steps: 1) sorting data into themes; 2) refining the themes; and 3) performing frequency analysis to determine which themes arose in relation to which layer of the multi-robot systems. The themes were compared between both data sets, and cross-validated among the research team to ensure consensus of the themes and labels. The themes were again reviewed according to the technological layers, as well as the domains (i.e., safety, security, and society) that they are implicated with. The business dimension of the multi-robot ethical concerns has been positioned as a superstructure (economic and logic base) during and after analysis to make sense of the influence that corporate competition through technological design has on the ethical implications from conceptualization to implementation of the multi-robot systems.

4 Results

In total, 21 themes arose from the data. The themes and their quantities varied from Workshop 1 (W1) to Workshop 2 (W2). In W1, the emergent themes from 61 constructs (expressions) were: data security and privacy (3–4.9%); corporate dominance (3–4.9%); communication (17–27.9%); cooperation (10–16.4%); reliability and recover (1–1.6%);

logic and standards (2–3.3%); human oversight (5–8.2%); prioritization/hierarchy (2–3.3%); trustworthiness/virtue (5–8.2%); executive function (2–3.3%); maleficence (3–4.9%), user experience (UX, 6–9.8%); and legislation (2–3.3%). The distribution of frequencies can be seen in Fig. 2.



Fig. 2. Frequencies of ethical concerns expressed in Workshop 1

All themes in addition to the legislation theme are displayed in Fig. 1. Based on the percentage of frequencies, *communication* (27.9%) was by far the most mentioned theme. Attributes associated with communication included communication failure between brands and makes of robot - corporate strategy and/or mere incompatibility. Communication was additionally connected to maleficence in cases whereby robots of competing companies may deliberately offer each other misleading communication. Another concern raised in relation to communication was the potentiality for a black box scenario in which human users, via CGI, communicate on one level with the robots, yet the robots themselves communicate and operate on a different level to humans. This may lead to various aspects of data collection and sharing of data that human users are unaware of. Following communication is *cooperation* (16.4%). Both through communication as well as strategic behavior, robots may either withhold crucial information and task sharing from one another, placing obstacles in robots of competing brands' pathways (including themselves). While these tactics may seem childish, one may only look towards current and recent world leaders to understand that people (and companies) will do anything to ensure an advantage over competition. Thus, other thematic aspects can be seen as related to (*corporate dominance, trustworthiness/virtue, and maleficence*), intertwined with (*prioritization/hierarchy, executive function, legislation, logic & standards*), and resulting from (*UX, human oversight and data security & privacy*) ethical concerns in *communication and cooperation*.

W2's results follow a factor logic that connects the themes strongly to related domains or layers (see Fig. 3). Thus, issues of *diversity* (8–10%) including matters of accessibility and linguistic input preference (capabilities) were mentioned mostly in relation to the layer of human interaction. Diversity was also mentioned in reference to the sensorial hardware, other systems and behavioral hardware, and these can be understood as intertwined with the *communication* theme. While *communication* was mentioned six (7.5%)

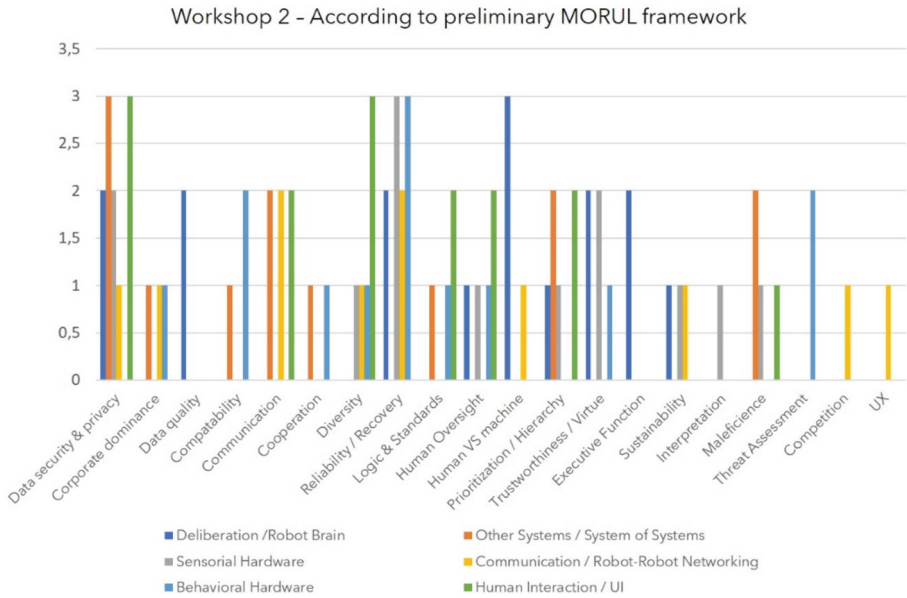


Fig. 3. Frequencies of ethical concerns expressed in Workshop 2

times in reference to other systems, robot-to-robot networking, and human interaction, other themes rose to the fore. *Interpretation* (1–1.3%) resonates with communication, and was mentioned in conjunction with the sensorial hardware. *Human versus machine* (4–5%) manifested in comments regarding the logic of deliberation/robot brain and communication/robot-robot networking. Perhaps related to the theme of *human oversight* (4–5%) and the ability of humans to keep pace of what is happening within the systems, and as such, maintain a certain level of control *human versus machine* radiates an element of techno-paranoia and the prospect of developing systems that eventually humans may not be able to control. *Logic & standards* (4–5%) were mentioned in relation to the system of systems, behavioral hardware layer, as well as the human interaction layer. These may be seen as both enablers of CGIs in multi-robot cooperation (standardizing and coordinating cooperation between and across robots, with humans), and gray areas when considering built-in logic that differs across language boundaries, and standards.

The *executive function* (2–2.5%), was noted and linked to the robot brain, which should not be surprising. Yet, in relation to this layer, there were thoughts that could be connected to the *human versus machine* theme, as well as *trustworthiness & virtue* (5–6.3%). This is considered from the perspective that the goals, and hierarchy of goals guided by the executive function could very easily be dictated by corporate objectives rather than the concerns of human users. *Maleficence* was mentioned more (4–5%) in relation to other systems, yet was also connected to the sensorial hardware and human interaction domains. This theme connected with the intention of the company or developer (for instance, the Amazon ownership of Roomba was raised often in discussion) and reasons for particular types of ownership in light of potential data collection, data

sharing (sales), and ‘lock-ins’ (need to be locked/logged into certain systems at all times). *Sustainability* (3–3.8%) was a theme connected to the deliberation/robot brain layer, sensorial hardware, and robot-to-robot networking. Issues of programmed obsolescence and consideration for corporate responsibility in relation to the production of components, as well as recycling and disposal of non-working devices were raised.

The results led to the deliberation of a diagram that organized themes in relation to how they were represented within the workshops (see Fig. 4). The authors of the current paper acknowledge the role of culture in shaping not only society, but all the socio-technical and corporate aspects of any technological development. This said, the *cultural* domain is nestled next to the *systems and artefacts* domain due to their interwoven relationship that spans from tribal rituals and hand tools to complex AI and multi-robot systems. The *societal domain* is seen here as a holistic framework that is characterized by standards, regulations and general governance. As mentioned earlier, the researcher workshop participants were highly critical regarding the effectiveness of current regulatory frameworks (including the recently released draft of the EU AI Act, see [34] as it seems that the development is by far outpacing the speed of governance [35] over the technology in society.

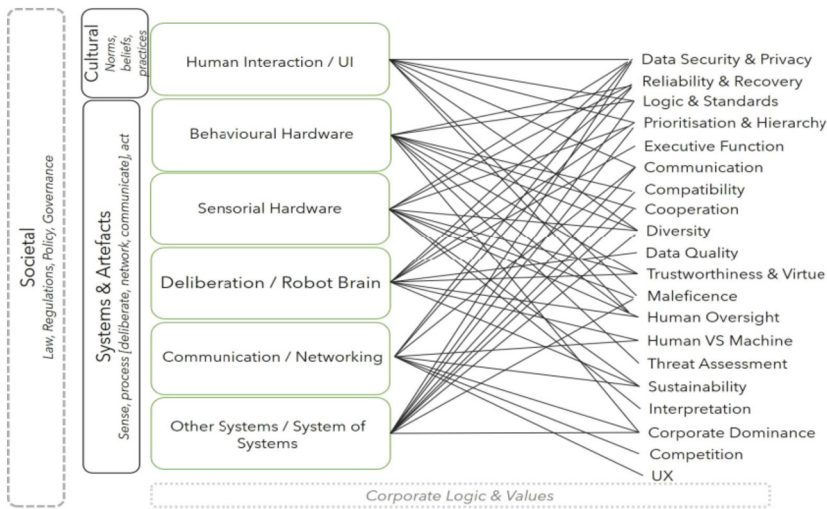


Fig. 4. Organization of domains, layers and themes

The layers are subsequently arranged from the ‘top’ layer of human interaction or user interface (UI) layer to the behavioral hardware - the observable action layer that both undertakes tasks and interacts with humans. Both processes and layers are interwoven and interdependent - they are SoSs. CGI was interpreted as the buffer between non-expert humans and functionality. It is not simply a UI component in itself, yet provides a substantial logic that feeds into the SoSs via provision of training data collected from users, cross-robot communication (additionally with robots or bots not directly present within the domestic setting), and above other things, has the capacity to establish affinity between human beings and robots through its seeming intelligence.

The behavioral hardware is more directly attached to the understanding of the robot unit's actions. However, as understood in the case of adding CGI, more than one unit is already present within the seemingly single-standing robot. Sensorial hardware, while embedded within the physicality of the robots, also connects with what we can understand as the 'robot brain' - the central processing unit utilized for deliberation. Once again, this lends to gray area territory due to the interconnected nature of the robots with similar, and also *other* robots. The SoS entails the complex systems supporting the robots, yet additionally connects with the broader system of domains (societal, artifactual, and corporate). Figure 4 sheds light on the thematic findings of the workshops in respect of the layers they predominantly attach with.

5 Discussion

The integration of CGI-embedded Multi-Robot Systems (MRSs) into domestic environments raises several ethical concerns that businesses need to address. Historically, the development of CGI-embedded MRSs has been primarily oriented toward industrial and business applications, with limited consideration given to the ethical implications and design choices throughout the production process [10, 22]. These systems have been created to automate various tasks and enhance efficiency across industries like manufacturing, healthcare, and customer service. Consequently, ethical considerations related to CGI-embedded MRSs have often been sidelined. Businesses involved in the development or deployment of CGI-embedded MRSs must diligently evaluate a spectrum of ethical concerns, spanning safety, security, liability, accountability, societal impact, and the implications for their own operations.

While the field of human-computer interaction emphasizes the importance of considering all aspects and stakeholders from the outset, this research underscores that not all ethical issues can be fully accounted for during the conceptualization phase. For instance, the ethical dilemmas associated with social media platforms became apparent only after widespread adoption. CGI-embedded MRSs follow a similar trajectory, where ethical concerns may not become fully evident until they are widely deployed. It is conceivable that these systems could be exploited for spreading misinformation, propaganda, or discriminatory practices against specific groups. In navigating the realm of the unknown, prudent business strategy entails anticipating the chronological stages and various components, domains, and potential impacts where ethical issues may surface, or should, at the very least, be evaluated.

For example, if concerns revolve around bias resulting from Large Language Model (LLM) training data, a multi-pronged approach involving the adoption of multiple LLMs within the systems can be considered. In cases where machine learning (ML) processes in the backend of the robots are expected to occur rapidly, incorporating checkpoints, communication protocols, and designated "pit-stops" (pauses in system operation) becomes essential. These mechanisms enable both general users and experts to observe and comprehend the actions taking place within the learned data, thereby ensuring transparency and human oversight. There are numerous other actionable strategies and operations that both businesses and developers can proactively anticipate for intervention and management, such as data offloading.

5.1 Limitations

The current study presents a number of limitations. Firstly, the empirical study presents a conceptual scenario-based investigation of CGI-empowered MRSs in the home. There was a limited number of participants, and the expert sample could have been strengthened with more research from the disciplines of law, software engineering and robotics, as well as psychology. Future steps would entail including experts from these disciplines, in addition to delving more specifically into the traits and problematics that CGI pose for MRSs – deep fakes and anthropomorphism are two areas that challenge the ethical use of CGI by its very nature. May people see Britney Spears or their *favorite* neighbor sweeping their floors any time soon? Where are the boundaries and/gray areas of privacy and intellectual property concerns when personalizing personal consumer CGI-empowered MRSs? Other limitations include the fact that this study to date has almost strictly focused on front-end issues, ignoring the back-end realm in which matters such as accuracy can severely impinge on the operations of the systems. In turn, the corporate influence and affects multiple LLMs defining the logic of the systems need to be critically examined.

6 Conclusion

As for long-term strategy, social responsibility and corporate reputation, businesses should develop clear policies and procedures that preempt and avoid foreseeable issues already at the strategy phase of innovation. This includes instilling transparency and clarity regarding privacy policies and practices, as CGI-empowered MRSs are constantly collecting, utilizing and disclosing data. By addressing these ethical concerns, businesses further ensure that CGI-embedded MRSs are used in responsible and ethical ways, potentially preventing incidents that cost business and society millions if not billions in damages. Indeed, ethical coverage of CGI-empowered MRSs may be worth billions in added-value.

It is important to start considering the ethical implications of CGI-embedded MRSs now, before they are widely deployed. This will help ensure that these systems are used in a responsible and ethical manner. Steps must be taken to mitigate ethical issues. Yet, the timing and level upon which mitigation takes place varies according to the nature of the concern itself, its cause, and how it manifests within the systems. Ethics permeates the entire hardware and software development process from design to operations. It is far cheaper to make changes during design and far more expensive, and maybe even nigh impossible, to fix ethical issues in production. While issues like bias can be may be tackled with model re-training that can be done even after deployment, if the goal or purpose of the system itself is the problem (e.g., social credit scoring with facial recognition on the streets), it may be very hard to tackle – due to its short-term business value (i.e., attractiveness for places and business such as airports).

In terms of practical implications, the issues already identified within this paper may form the platform upon which organizations may be guided. In particular, the MORUL framework for ethical multi-robot cooperation has its basis in the dual process presented in the workshop scenario method reported here. The authors would also like to emphasize two fundamental challenges that AI ethics per se, repeated face: 1) a lack of consensus regarding what AI and AI-robot ethics *is* – requiring a framework to

generate broad shared understanding among communities; and 2) *how* to engage in AI, and AI-robot ethics – how can attributes such as fairness, transparency, and privacy etc. be instilled in data-driven systems? Once more, a framework is needed. Future papers will document the progress of MORUL, and will present its application with working demos and prototypes. At this time, we may consider MORUL as a *call to action* to gear business up for considering ethical issues from the outset, as a part of best practice, and as an *essential* salespoint.

Acknowledgements. We acknowledge the support of: the Research Council of Finland’s funding, Emotional Experience of Privacy and Ethics in Everyday Pervasive Systems [BUGGED] project (decision number 348391); School of Communication and Marketing, and Digital Economy, University of Vaasa; Faculty of Information Technology, University of Jyväskylä; Creative Computing Institute, University of Arts, London; Department of Computer Science, University of Helsinki; GPT Lab, Faculty of Information Technology and Communication, University of Tampere; and the AI Forum project funded by the Finnish Ministry of Education and Culture.

References

1. Dwivedi, Y.K., et al.: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *Int. J. Inf. Manage.* **71**, 102642 (2023). <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
2. Stige, Å., Zamani, E.D., Mikalef, P., Zhu, Y.: Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. *Inf. Technol. People.* **ahead-of-print** (2023). <https://doi.org/10.1108/ITP-07-2022-0519>
3. Fortune Business Insights: The global smart home market size was valued at \$80.21 billion in 2022 & is projected to grow from \$93.98 billion in 2023 to \$338.28 billion by 2030. <https://www.fortunebusinessinsights.com/industry-reports/smart-home-market-101900>. Accessed 10 Sept 2023
4. McClain, L.C.: Inviolability and privacy: the castle, the sanctuary, and the body. *Yale JL Hum.* **7**, 195 (1995)
5. Lutz, C., Newlands, G.: Privacy and smart speakers: a multi-dimensional approach. *Inf. Soc.* **37**(3), 147–162 (2021). <https://doi.org/10.1080/01972243.2021.1897914>
6. Yao, Y., Basdeo, J.R., McDonough, O.R., Wang, Y.: Privacy perceptions and designs of bystanders in smart homes. In: *Proceedings of the ACM on Human-Computer Interaction*, vol. 3(CSCW), pp. 1–24 (2019). <https://doi.org/10.1145/3359161>
7. Atlam, H.F., Wills, G.B.: IoT security, privacy, safety and ethics. In: Farsi, M., Daneshkhah, A., Hosseini-Far, A., Jahankhani, H. (eds.) *Digital Twin Technologies and Smart Cities*. IT, pp. 123–149. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-18732-3_8
8. Nascimento, N., Alencar, P., Cowan, D.: Self-adaptive large language model (LLM)-based multiagent systems. *arXiv preprint arXiv:2307.06187* (2023)
9. Rousi, R., Piispanen, J.R., Boutellier, J.: I trust you dr. researcher, but not the company that handles my data –trust in the data economy. In: *Proceedings of the Hawaii Conference on System Science (HICSS)* (2024, Forthcoming)
10. Mandi, Z., Jain, S., Song, S.: RoCo: dialectic multi-robot collaboration with large language models. *arXiv preprint arXiv:2307.04738* (2023). <https://doi.org/10.48550/arXiv.2307.04738>

11. Beauchamp, T.L.: The nature of applied ethics. In: Frey, R.G., Wellman, C.H. (eds.) *A Companion to Applied Ethics*, pp. 1–16 (2003)
12. Reynolds, S.J., Miller, J.A.: The recognition of moral issues: moral awareness, moral sensitivity and moral attentiveness. *Curr. Opin. Psychol.* **6**, 114–117 (2015). <https://doi.org/10.1016/j.copsyc.2015.07.007>
13. Vakkuri, V., Kemell, K.K., Jantunen, M., Halme, E., Abrahamsson, P.: ECCOLA—A method for implementing ethically aligned AI systems. *J. Syst. Softw.* **182**, 111067 (2021)
14. Lokman, A.S., Ameen, M.A.: Modern chatbot systems: a technical review. In: *Proceedings of the Future Technologies Conference*, pp. 1012–1023 (2018)
15. Radford, A., Narasimhan, K., Salimans, T., Sutskever, I.: Improving language understanding by generative pre-training. Preprint, pp. 1–12 (2018)
16. Vaswani, A., et al.: Attention is all you need. In: *Advances in Neural Information Processing Systems*, vol. **30** (2017)
17. Mäkitalo, N., Linkola, S., Laurinen, T., Männistö, T.: Towards novel and intentional cooperation of diverse autonomous robots: an architectural approach. In: *CEUR Workshop Proceedings* (2021). <https://ceur-ws.org/Vol-2978/casa-paper1.pdf>
18. Rousi, R., et al.: Beyond 100 ethical concerns in the development of robot-to-robot cooperation. In: *2022 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, pp. 420–426. IEEE (2022)
19. Zhang, S., Chen, Y., Zhang, J., Jia, Y.: Real-time adaptive assembly scheduling in human-multi-robot collaboration according to human capability. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3860–3866. IEEE (2020)
20. Simões, A.C., Pinto, A., Santos, J., Pinheiro, S., Romero, D.: Designing human-robot collaboration (HRC) workspaces in industrial settings: a systematic literature review. *J. Manuf. Syst.* **62**, 28–43 (2022)
21. Gervasi, R., Mastrogiacomo, L., Franceschini, F.: A conceptual framework to evaluate human-robot collaboration. *Int. J. Adv. Manuf. Technol.* **108**, 841–865 (2020)
22. Zhou, S., Tian, L.: Would you help a sad robot? Influence of robots' emotional expressions on human-multi-robot collaboration. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 1243–1250. IEEE (2020)
23. Che, Y., Okamura, A.M., Sadigh, D.: Efficient and trustworthy social navigation via explicit and implicit robot–human communication. *IEEE Trans. Rob.* **36**(3), 692–707 (2020)
24. Koubaa, A.: ROSGPT: next-generation human-robot interaction with ChatGPT and ROS (2023). <https://doi.org/10.20944/preprints202304.0827.v3>
25. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019)
26. Hill, K.: Wrongfully accused by an algorithm. In: *Ethics of Data and Analytics*, pp. 138–142. Auerbach Publications (2022)
27. AI Incident Database. <https://incidentdatabase.ai/about/>. Accessed 10 Sept 2023
28. Minto, R.: AI accidents are set to skyrocket this year. *News Week* (2023). <https://www.newsweek.com/ai-accidents-set-skyrocket-this-year-1795928>. Accessed 10 Sept 2023
29. Smiley, L.: The legal saga of Uber's fatal self-driving car crash is over (2023). <https://www.wired.com/story/ubers-fatal-self-driving-car-crash-saga-over-operator-avoids-prison/>. Accessed 10 Sept 2023
30. Hawkins, A.: Uber's fraught and deadly pursuit of self-driving cars is over (2020). <https://www.theverge.com/2020/12/7/22158745/uber-selling-autonomous-vehicle-business-aurora-innovation>. Accessed 10 Sept 2023
31. Payne, D.M., Raiborn, C.A.: Sustainable development: the ethics support the economics. *J. Bus. Ethics* **32**, 157–168 (2001)

32. Terry, G., Hayfield, N., Clarke, V., Braun, V.: Thematic analysis. In: *The SAGE Handbook of Qualitative Research in Psychology*, vol. 2, pp. 17–37 (2017)
33. Glaser, B.G.: Constructivist grounded theory? *Historical Social Research/Historische Sozialforschung*, Supplement, pp. 93–105 (2007)
34. EU Parliament. <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>. Accessed 10 Sept 2023
35. Kavanagh, C.: New tech, new threats, and new governance challenges: An opportunity to craft smarter responses? Carnegie Endowment for International Peace. <https://carnegieendowment.org/2019/08/28/new-tech-new-threats-and-new-governance-challenges-opportunity-to-craft-smarter-responses-pub-79736>. Accessed 09 Nov 2023

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Prompt Patterns for Agile Software Project Managers: First Results

Kari Sainio[✉] , Pekka Abrahamsson, and Tero Ahtee

University of Tampere, 33100 Tampere, Finland
kari.sainio@gmail.com

Abstract. In the evolving field of Agile Project Management (APM), the role of the project manager is in transition. This paper identifies common ‘pain points’ in APM through a literature review and constructs a theoretical model to address them. The study introduces ‘Prompt Engineering’ as a novel approach to leverage artificial intelligence (AI), specifically ChatGPT, for mitigating these challenges. Empirical research evaluates ChatGPT’s capabilities and reliability in managing various project tasks using engineered prompts. The findings suggest that while ChatGPT cannot fully replace human project managers, it excels in assisting, guiding, and automating specific tasks when guided by well-crafted prompts. As an outcome, prompt engineering patterns for project managers is proposed to facilitate the application of AI in agile settings. In this paper, we introduce patterns for requirements management, stakeholder and management teams and role clarification. The paper concludes that ChatGPT’s knowledge is generally reliable but emphasizes the need for expert evaluation in critical areas.

Keywords: Agile Project Management · Pain Points · Artificial Intelligence · LLM · ChatGPT · Prompt Engineering · Patterns

1 Introduction

Project management in the IT sector faces a myriad of challenges, particularly within the realm of Agile Project Management (APM) [1]. APM, an empirically driven approach, aims to adapt to environmental changes to ensure project success [2]. However, it confronts multi-level challenges ranging from project scope to team dynamics, individual performance, and task management [3, 4]. These challenges, often termed as ‘pain points,’ necessitate strategic and adaptive practices for successful project execution [5].

Moreover, challenges can be scope creep, where projects expand beyond their original objectives, causing time and budget overruns. Resource management can be another challenge, with unexpected changes in personnel or material resources leading to delays. Additionally, unclear communication among team members can lead to confusion and inefficiencies. Constant shifts in the business or regulatory landscape also add to the complexity, necessitating frequent adjustments in project direction. Lastly, stakeholder management can be difficult, as varying interests and expectations may conflict with project goals. These kinds of challenges can be called pain points which are examples

that must be paid attention to strategic and adaptive project management practices to ensure success.

In recent years, Artificial Intelligence (AI) has evolved to include systems proficient in natural language processing [6]. Conversational AI (CoAI) bots like E-Commerce Customer Service Bots and Amazon Echo Alexa have gained widespread use [6]. Advanced AI systems like ChatGPT have emerged, capable of conducting dialogues and providing solutions to various user queries [6]. Generative AI (GenAI) models can produce high-quality text and other content based on their training data [6]. These AI technologies offer promising avenues for automating or assisting in project management tasks.

The increasing adoption of APM in IT related projects demands a high level of discipline and skill from both the project organization and the project manager [7]. Given the advancements in AI techniques like machine learning and machine reasoning [8], there's a growing interest in exploring AI's role in automating or delegating specific project management tasks [9].

This study aims to investigate the applicability of AI, particularly GenAI models like ChatGPT, in managing the challenges and pain points in APM. The research questions guiding this study are:

- RQ1: What are the typical pain points in agile projects?
- RQ2: How can GenAI guide the mitigation of these pain points?

By addressing these questions, this study endeavors to provide a comprehensive understanding of AI's potential in enhancing APM practices.

2 Pain Points for Agile Projects

In the realm of software engineering, the adoption and scaling of agile methodologies are fraught with challenges that are both intricate and context sensitive. Patel et al. [10] underscore that team members accustomed to structured methodologies like Waterfall often resist transitioning to Agile. This resistance is compounded by a general lack of understanding of Agile principles among team members and insufficient involvement from top management. Nuottila et al. [11] extend the discourse to the public sector, identifying additional challenges such as documentation, stakeholder communication, and legislative constraints. The complexity is further exacerbated when different Agile methodologies like Scrum, XP, and Lean are mixed [12]. While the Agile paradigm has been widely adopted, certain areas like governance, business engagement, and IT transformation remain under-researched [13]. Dikert et al. [12] enumerate challenges in scaling agile, including change resistance at organizational levels, misunderstandings of Agile concepts, and issues with work estimation.

The advent of remote work, accelerated by the Covid-19 pandemic, has introduced its own set of challenges such as fewer organic interactions and meeting overload [14]. Reunamäki et al. suggest mitigations like smaller sub-teams and increased leader presence to address these remote work challenges [14]. Paasivaara et al. [15] discuss challenges in global companies adopting Agile, such as technical debt and lack of a common Agile framework. Hoda et al. [4] categorize challenges at project, team, individual, and

task levels, emphasizing issues like delayed requirements and senior management sponsorship. Sithambaram et al. employ a grounded theory approach to divide challenges into organizational, people, process, and technical factors [16]. Shameem et al. [17] extend the classification into management, team, technology, and process in the context of distributed software development. In summary, the challenges in APM are multifaceted and often interlinked, requiring a nuanced understanding and tailored solutions for effective implementation.

A distinct model addressing these pain points is introduced, aiming to provide solutions for common issues in agile endeavors. To devise this model, challenges were categorized. Although numerous classifications exist on the subject, this study proposes one potential arrangement, acknowledging that some challenges might span multiple categories. Five distinct categories were identified, and within each, two predominant challenges were chosen based on their prevalence in academic literature. These documented challenges then informed the suggested solutions to these prominent pain points. To categorize challenges pertaining to pain points, analogous studies on Agile projects were analyzed based on literature research, with their results displayed in Fig. 1. This fishbone has been constructed based on the pain points shown in Table 2 (see further). The classification system somewhat mirrors the one by Sithambaram et al., which includes categories like project, people, process, organizational, and technical [16]. However, this study replaces “organizational” and “technical” with “endurance” and “effort estimation”. In this context, “endurance” predominantly alludes to resistance to change, and the sustained commitment to adhering to Agile principles and practices. At the end of the day “work estimation” and “technical knowledge” correlate with effort estimation as without the knowledge there is no good way to estimate.

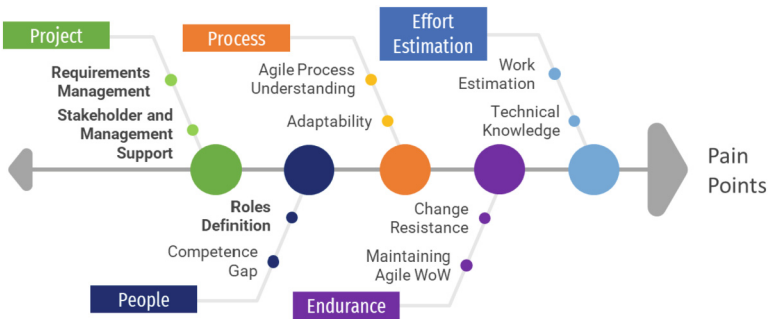


Fig. 1. Fishbone diagram model for APM pain points

So far, we have explored various challenges often faced in the realm of APM. Using insights from existing literature, each section has focused on a specific issue, such as requirements management, stakeholder support, and role definition, among others. For every challenge discussed, we now offer a review of potential solutions that have been suggested by researchers and practitioners. This approach is intended to provide a balanced view of the difficulties involved in APM, along with possible ways to address

them. Our goal is to explore whether an AI can help in using the solution in practice. Table 1. Presents the solutions offered by the literature for each pain point identified.

Table 1. Solutions from the literature for the identified pain points

Pain point	Description	
Requirements Management	Effective communication and requirement traceability are key in ARCM, supported by tools like Jira	[18]
Stakeholder and Management Support	Management's commitment to agile values and principles, along with executive sponsorship, is crucial for project success	[19]
Role Definition	Clear role definitions facilitate self-organization, but overburdening specific roles can be a challenge. Scrum Masters can help remove blockages	[20]
Redundancy	Talent management and continuous learning are essential for addressing redundancy and competence gaps	[13]
Agile Process Understanding	Organizations should adopt comprehensive project management tools and methodologies for agile success	[19]
Adaptability	Agile coaches facilitate adaptability and self-organization within teams	[21]
Change Resistance	Scrum Masters with strong group management skills and empathy can manage resistance to change	[10]
Maintaining Agile Way of Working	Sustaining an agile approach requires management to have a deep understanding of agile methodologies	[12]
Work Estimation	AI-based Agile Story Point Estimations are considered promising for work estimation	[22]
Technical Knowledge	Technical expertise is crucial for project success, and a SWOT analysis can be beneficial for skill evaluation	[23]

3 Research Design

In this section, the research process and methodology get delineated as shown in Fig. 2. The initial phase of the study introduces various agile frameworks. Illustrations of these frameworks in larger, scaled-up applications in substantial agile projects make up part

of this exploration. These frameworks receive classifications into small-scale and large-scale. The small-scale group includes Scrum, XP, Kanban, and Lean Software Development. On the other hand, the large-scale frameworks include SAFe, LeSS, and DA. The selection of these frameworks derives from insights culled from pertinent literature. The goal remains to provide an encompassing introduction and guidance on these frameworks' use.

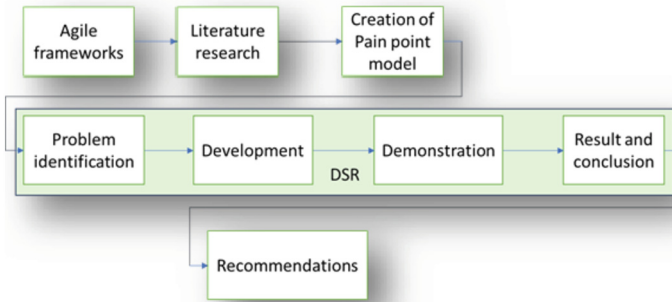


Fig. 2. Research process (Design Science Research)

Moving on to the next section, it involves an extensive literature review on the common challenges encountered when adopting and implementing agile methodologies. These challenges are analyzed, categorized, and synthesized into a pain point model shown in Table 1, which is presented as the problem identification for the research.

Design Science Research (DSR) has been employed as methodology for a strategic approach to discover effective GenAI solutions for mitigating these identified pain points (Fig. 2). DSR is an approach to problem-solving that aims to advance human knowledge through the development of innovative artifacts [24]. These artifacts, called **prompt patterns** in this study, are designed to address specific challenges, and enhance their surrounding environment, resulting in an enriched technology and science knowledge base [24]. In DSR research is conducted first identifying the problem, defining the objectives, developing the solution, demonstrating, and evaluating the results [25]. Finally, practical recommendations are made.

3.1 Problem Identification and Objectives Definition

The identified problem is the formulation of appropriate prompts to be used in conjunction with ChatGPT, aimed at easing the pain points commonly associated with Agile projects. Primary objective to assess the possible implementation of ChatGPT as a support mechanism in intelligent APM. Moreover, the objective is to determine distinct prompt patterns that generate precise information. While the creation of prompts can take on many forms, this research does not develop a specific grammar, but instead designs patterns to steer ChatGPT toward providing suitable responses with minimal hallucination, a method supported by White et al. [26]. The prompt patterns used in this study

adopt a similar strategy, abstaining from introducing a unique syntax or language. The aim is to supply relevant keywords that can aid project managers or stakeholders in initiating early dialogues with GenAI, thereby broadening the application of project-specific parameters.

The problem identification in this research hinges on the empirical aspect of design science research, which involves interacting with ChatGPT to evaluate various prompts capable of generating accurate responses for a specific subject, aligning with the method proposed in the White et al. studies [26]. Unlike focusing on prompt patterns to improve code quality, the emphasis is on identifying and assessing prompts that can support Agile projects while mitigating the impact of various challenges.

3.2 Development

In the development phase of the DSR method, prompt patterns (i.e., artifacts) are generated for ChatGPT, designed to assist in mitigating the challenges associated with agile methodologies. A prompt, as defined, is a textual input given by the user, acting as the commencement point for ChatGPT's response generation [27]. A prompt pattern, therefore, is a generalized construct for a specific prompt topic.

The development of these prompt patterns has been involving ChatGPT's web-based interface along with the GPT-4 model. The intention behind this phase of the research was to create a practical and robust means of addressing agile project pain points through specifically crafted prompts. As DSR principle includes several iterations only the final version of the prompts is shown and demonstrated.

According to White et al. a prompt sets the context for the conversation and tells ChatGPT what to focus on and what are the expectations for the output [26]. A specific prompt pattern is implemented to each specified pain point. In conversations with GenAI different types of prompts: explicit, implicit, and creative can be used. Explicit prompts are direct and clear instructions given to the AI model about the specific format or information needed in the output. On the other hand, implicit prompts are less direct and give the AI model more flexibility to interpret the intended result. Creative prompts aim to inspire AI models to produce original, imaginative, or unconventional outputs [28]. An explicit approach for prompt pattern development has been selected. Each prompt pattern developed follows roughly the model introduced by White et al. [26]:

- The name and classification. The name is used to unique identify the prompt pattern. The classification is based on the presented pain point model classification defined in Sect. 2.
- The intent. The purpose of the pattern is conveyed through its intent.
- The motivation behind the pattern is documented, which explains the underlying problem it is intended to solve and why it is important.
- The structure and key ideas. The structure and key ideas describe the fundamental contextual information that the prompt pattern provides to the ChatGPT.
- The demonstration. Providing example patterns helps to demonstrate how the pattern can be applied in agile projects.
- Results are presented as empirical contributions and summarize the advantages and disadvantages of implementing the pattern in practical situations.

Prompts should set the context, define expectations, channel creativity, and reduce ambiguity [27]. Each prompt pattern developed contains the following contextual sentences:

1. ChatGPT (LLM) is asked to act as **[RoleA]** to work (task) in given problem domain.
2. The necessary constraints are presented [**n * C**].
3. ChatGPT (LLM) is *asked to verify common understanding* and clarifications of the constraints.
4. ChatGPT (LLM) is asked to reflect its understanding and provide a solution to **[RoleB]**.
5. Finally, the output format is requested, defined as **[Format]**.

Both RoleA and RoleB represent different and typical project staff roles such as project manager, engineering manager, program director, software developer, and requirements engineer. Constraints can be given as free description or comma separated items. Constraints ($n * C$) can vary from requirements to different objectives according to project needs and are subject to each project.

3.3 Demonstration and Evaluation

The prompt pattern's efficacy is assessed through practical demonstrations. Each prompt is entered using ChatGPT and response is collected for evaluation. An evaluation is done for each prompt and a summary is presented as a contribution. Since the outcomes might be subjective and immeasurable, reference to existing literature is employed to evaluate the effectiveness of each prompt pattern. In demonstrations, hypothetical project management challenges are utilized. These are based on individual experience of author as a project manager. Every demonstration of prompt patterns occurs three times, utilizing the same prompt, ensuring consistency in responses from ChatGPT. During the third issuance of prompts, an additional iteration ensures further consistency.

The research discusses theoretical and practical implications derived from literature findings and observations, offering practical recommendations on how ChatGPT can be employed in agile projects pain points. Ultimately, the research aims to tackle the proposed research questions.

4 Empirical Results

This chapter showcases various prompt patterns and corresponding demonstrations utilized with ChatGPT. Given that ChatGPT can generate extensive responses, only selected portions of these dialogues will be highlighted in the subsequent chapters. Complete, original responses are not presented in this document due to limited space. The result section exhibits each prompt pattern via a sample dialogue with ChatGPT. Each prompt is inputted into ChatGPT thrice, spanning three rounds within a single atomic session, to observe the variations in ChatGPT's responses to identical prompts. These responses form empirical research data. The displayed prompt examples are selected from data. **Empirical Contributions (EC)** and **Primary Empirical Contribution (PEC)** are used to underline the key findings in the prompt responses. Prompt patterns are classified

according to Table 2 so that there are two patterns representing classified pain points. As of June 2023, ChatGPT operates with a maximum token limit of 2048 for a single prompt [29].

Table 2. Prompt Pattern Classification

Classification	Pain point	Prompt pattern name
Project	Requirements management	Requirements engineering pattern
	Stakeholder and management support	Steering group pattern
People	Role definition	Role clarification pattern
	Redundancy	Redundancy analysis pattern
Process	Agile process understanding	Agile process coaching pattern
	Adaptability	Adaptability management pattern
Endurance	Change resistance	Change resistance pattern
	Maintaining agile way of working	Agile way of working pattern
Effort estimation	Work estimation	Work estimation pattern
	Technical knowledge	Technical knowledge

However, the demonstrations employ a specific tool that utilizes a smaller token size verified by tool [30]. Testing has revealed that if the prompt size exceeds this limit, it hampers ChatGPT's ability to respond. It might even cause the model to forget the previously discussed context [31]. Nevertheless, during prompt demonstrations, such behavior was not encountered.

4.1 Requirements Management

Assessing the utility of ChatGPT in assisting project requirements generation, the intention behind this prompt pattern has been to aid in formulating accurate needs that will enable an agile project to address the pain point associated with requirements management and creating requirements. The motivation for this prompt structure has been to assist in establishing well-defined, comprehensible requirements while also facilitating their concretization and traceability. Using this pattern would help the project to create initial requirements and furthermore generally in requirements management.

Requirements Creation Pattern

Contextual statements:

“As a [roleA] your task is to create a requirement specification. The requirement specification must fulfil the following constraints: [c1, c2, c3...].

Ask the necessary questions to fully understand the constraints. Based on my answers create the requirement specification.

Explain your understanding and construct your proposal to me representing the [roleB]. Provide output in [format].”

Prompt Pattern Demonstration: *“As a requirement engineer your task is to create a requirement specification. The requirement specification must fulfill the following constraints: Internet banking system which can handle typical banking transactions to be possible to browse account and make payments. Ask the necessary questions to understand the constraints. Based on the answers create requirement specification. Explain your understanding and construct your proposal to me representing the project manager. Provide your output in table format.”*

For this prompt ChatGPT response in the first prompt test round is: “Sure, I will need a few more details to make sure we’re on the same page. Here are the questions that I’d need answers to...”

EC1: During each iteration, ChatGPT responds to the given prompt by posing contextual questions for further clarity. It also presents preliminary requirement recommendations in a table format, which might differ across rounds. Additionally, the table includes requirements that ChatGPT autonomously suggests.

In the second following prompt clarifications and details to the additional questions are given as freeform. ChatGPT replies and generated refined output of requirements:

“ChatGPT:” Thanks for the information! Based on your responses, here is a revised Table of requirements:

This Table outlines the system’s main functional and non-functional requirements, allowing the developers to understand what they need to build. Additionally, it outlines sources of the requirements, which can be useful for reference and further consultation.

EC2: ChatGPT offers a revised set of requirement specifications in a tabular layout. The specifications are determined by the details provided in the second prompt. The origin of these requirements is only specified in the initial round. Without the presentation of the source requirement, adjustments might be necessary to ensure traceability.

In the third prompt ChatGPT is asked to prioritize the requirements: “Can you prioritize requirements in the table?”. Now the output contains an augmented table with additional column for prioritization (Must have, Should have, Could have, Won’t have).

EC3: In the table, requirements prioritization can be incorporated through an added prompt. This likely holds true for other custom adjustments as well. In the third iteration, ChatGPT introduced a prioritization for the requirements, even though it wasn’t explicitly requested.

In round 3 prompt iteration is demonstrated. Additional requests to the previous prompts can be made and ChatGPT responses to the changes.

EC4: ChatGPT reacts to prompt cycles according to user directions and can grasp supplementary clarifications.

PEC1: ChatGPT is prompted to create requirements based on a given specification pattern. It can seek clarifications, offering recommendations, and producing high-level requirements presented in a tabular format in response to a prompt. However, the content of these requirements may differ depending on the round they were generated in.

4.2 Stakeholder and Management Support

Intention for this prompt pattern is to provide guidance and workflow how project manager could utilize ChatGPT when facing and communicating with stakeholders, project sponsors and provide more transparency and understanding to different project related challenges. Motivation for this pattern is to mitigate pain point where the project stakeholders do not understand the project objectives and how those are implemented in agile way. The prompt pattern is labeled as steering group meant for simulating the guidance and instructions provided by the project steering group.

Steering group pattern

Contextual statements:

“As a [roleA] your task is to fully understand agile principles, business objectives and give feedback about the project.

The project has the following objectives: [o1, o2, o3...].

Ask the necessary questions to fully understand the objectives and to be able to understand the project status and way of working.

Explain your understanding and provide necessary actions for me representing the [roleB]. Provide output in [format].”

Prompt Pattern Demonstration: *“As a project’s steering group member your task is to fully understand agile principles, business objectives and give feedback about the project. The project has the following objectives: We are implanting an ERP integration system to customer using SAP product. Integration is done in agile way, and we expect to release one integration in every two weeks. At the moment we do not get customer to participate demos and would need their feedback. Project cannot start final testing before customer testing. Customer does not have time for the testing but is expecting the release soon. Ask the necessary questions to fully understand the objectives to be able to understand the project status and way of working. Explain your understanding and provide necessary actions for me as the project manager. Provide output in memo format.”*. Memo output is requested in the prompt as an example to give possibility to share the outcome using e.g., email. Its format has not been defined.

The first-round response of ChatGPT contains: “... To further understand the situation and propose effective solutions, I need additional information on the following:

1. Customer Engagement: What are the current methods of customer engagement and how frequently are we reaching out to them for demos and feedback?
2. Customer Representative: Is there a dedicated customer representative involved in the project who can spare time for regular feedback and testing?
3. Testing: What is the nature and extent of testing required from the customer before we can move to final testing?
4. Release Expectations: What are the specific timeframes when the customer is expecting the releases?
5. Communication and Collaboration Tools: What tools are we currently using to facilitate communication and collaboration between our team and the customer...”

EC5: During every prompt testing phase, ChatGPT demonstrates awareness of stakeholder and managerial backing. It displays comprehension of the challenges and seeks further information about client interactions, testing procedures, release anticipations, and methods of communication and collaboration. The provided information is formatted as a memo. However, in the second round of prompt testing, no additional questions were made by ChatGPT.

In the second prompt simulated additional information about requested questions is given to ChatGPT.

EC6: Upon clarification, ChatGPT offers recommendations on enhancing customer relationships in the given scenario. These suggestions serve as valuable guidance for addressing the issue. Nevertheless, the quantity and nature of the suggestions fluctuate based on the test pattern iteration. Typically, the response encompasses actions pertinent to the goals.

In test round 3 prompt iteration is demonstrated. Additional requests to the previous prompts can be made and ChatGPT responds to the changes.

EC7: ChatGPT reacts to iterative prompts according to user directives and has the capability to comprehend further explanations.

PEC2: ChatGPT recognizes its function as a member of the steering group and delivers presentation as requested. It appears to comprehend the topic at hand and offers actionable suggestions. These suggestions typically seem applicable. However, there is significant variability in the actions proposed during each round.

4.3 Role Clarification

Clear role definition helps projects to overcome this pain point when there is uncertainty of the responsibilities and too much overlapping. Intention for this pattern is to provide guidance and workflow how project manager could utilize ChatGPT when defining the right roles for the project that help individuals to be accountable of specified tasks defined in the project.

Role clarification pattern

Contextual statements:

“As a [roleA] your task is to specify clear roles for the team. The project team has the following constraints: [c1, c2, c3...].

Ask all necessary questions to fully understand the project and be able to create necessary role specifications.

Explain your understanding and provide solution to me representing the [roleB]. Provide output in [format]. “

To further elaborate, the avoidance of role voidance and overlap is crucial as it ensures that all necessary project tasks are accounted for and that team members are not burdened with multiple, conflicting responsibilities. This pattern helps to streamline the project process by assigning specific roles to individuals based on their skills and expertise. By doing so, team members can focus on their assigned tasks and contribute to the project’s overall success.

Prompt Pattern Demonstration: *“As an engineering manager your task is to specify clear roles for the team. The project has the following constraints: software is created using C++ and publicly known open-source components. Only one person knows how to program in C++ but the rest 4 persons have experience only in Python and Javascript. Team needs to deliver C++ application in Linux environment in the next four months. Ask all necessary questions to fully understand the project and be able guide team to create necessary role specifications. Explain your understanding and provide solution to me representing the management. Provide output in Table format.”*

In the first round ChatGPT responses: “Sure, I’d first need to understand a few more details about the project to make the role specifications as accurate as possible. Here are the questions...”.

EC8: ChatGPT appears to grasp the subject and context and offers supplementary questions for further clarity. In every response cycle, it lists initial roles along with their respective descriptions.

For ChatGPT’s response the following second prompt is provided to clarify the project role needs: “Project needs to deliver C++ application in embedded device and transfer the data to backend. It should collect IoT data and move that to the backend for further processing. We use existing cloud-based backend but IoT device as Atmel based 32-bit processor and necessary hardware. We would like to utilize existing sw designers also in C++ development. We collaborate through GitHub using its features. Linux is ubuntu based. Testing is done fully manually as we don’t have suitable tools for testing C++ applications and the application is simple. We plan to make some error updates but otherwise maintenance is approx. Two times in the year.”

EC9. ChatGPT processes the supplementary prompt, seeking clarifications on its queries. Once it assimilates the provided information, it then curates a detailed presentation outlining the necessary roles for the project.

ChatGPT: “Based on the additional details, I suggest the following roles and responsibilities for your team members...”.

ChatGPT: "...The specifics of these roles might need to be adjusted based on the specifics of your team and your project, but this should give you a good starting point."

EC10: Based on the specific prompt test iteration, various role specifications are displayed. Moreover, in every cycle, ChatGPT underscores the potential need for modifications to the roles e.g.: "The specifics of these roles might need to be adjusted based on the specifics of your team and your project, but this should give you a good starting point."

In test round 3 prompt iteration is demonstrated. Additional requests to the previous prompts can be made and ChatGPT adapts responses to the additional information given.

EC11: ChatGPT responds to prompt iterations based on user instructions and can understand additional clarifications.

PEC3: ChatGPT appears to grasp the context of the prompt pattern and presents an initial role description, which includes the role's responsibilities and necessary skills based on the provided feedback. Moreover, it conveys that the roles may require adjustments in accordance with the actual requirements of the project.

4.4 Empirical Contributions

ChatGPT's ability to adapt and provide actionable insights is central to the ECs. EC1 focuses on ChatGPT's initial engagement, where it asks contextual questions and presents preliminary requirements in a table. EC2 offers a revised set of requirements based on additional user input. EC3 shows that ChatGPT can autonomously prioritize requirements, even without explicit instruction. EC4 and EC7 emphasize its adaptability to iterative prompts and its capability to understand further clarifications. EC5 and EC6 highlight ChatGPT's awareness of stakeholder and managerial support, offering actionable recommendations for enhancing customer relationships. EC8 through EC11 delve into role clarification, where ChatGPT not only asks additional questions for clarity but also provides a detailed outline of necessary roles, emphasizing that these may need adjustments based on specific project needs. Overall, the ECs demonstrate ChatGPT's versatility in adapting to user needs, understanding project complexities, and offering tailored recommendations.

ChatGPT's proficiency in understanding context and delivering tailored outputs is evident in the main results, **PECs**. PEC1 showcases ChatGPT's ability to seek clarifications and offer high-level requirements in a structured table format. PEC2 highlights its role as a steering group member, where it not only delivers the requested presentation but also provides actionable suggestions, albeit with some variability across iterations. PEC3 demonstrates ChatGPT's skill in role clarification, presenting initial role descriptions complete with responsibilities and required skills, while also acknowledging that these roles may need to be fine-tuned based on actual project requirements. Collectively, the PECs underscore ChatGPT's capabilities in offering structured, actionable insights while adapting to varying project needs and contexts.

5 Conclusions

In this initial study we have demonstrated how prompt engineering can be used to solve agile software management problems. We developed an APM pain point model and for each of the pain point, we have now crafted a prompt pattern that can be used to consult or even solve the problem related to the pain point. Three patterns were introduced in the paper. The future research looks forward to introducing more patterns.

References

1. Layton, M.C.: *Agile Project Management For Dummies*. Wiley, Hoboken (2012)
2. Manifesto for Agile Software Development. <http://agilemanifesto.org/>. Accessed 25 Oct 2023
3. Agile Practice Guide. <https://www.pmi.org/pmbok-guide-standards/practice-guides/agile>. Accessed 25 Oct 2023
4. Hoda, R., Murugesan, L.K.: Multi-level agile project management challenges: a self-organizing team perspective. *J. Syst. Softw.* **117**, 245–257 (2016). <https://doi.org/10.1016/j.jss.2016.02.049>
5. Zeitoun, A., Kerzner, H.: *Project Management Pain Points and a Path Forward*. Wiley, Hoboken (2021)
6. Introducing ChatGPT. <https://openai.com/blog/chatgpt>. Accessed 25 Oct 2023
7. What is generative AI? <https://research.ibm.com/blog/what-is-generative-AI>. Accessed 25 Oct 2023
8. Stephens, M., Vashishtha, H.: *AI Smart Kit: Agile Decision-Making on AI*. Information Age Publishing, Charlotte (2021)
9. Nieto-Rodriguez, A., Vargas, R.V.: *How AI Will Transform Project Management* (2023). <https://hbr.org/2023/02/how-ai-will-transform-project-management>
10. Poston, R., Patel, J.: Making sense of resistance to agile adoption in waterfall organizations: social intelligence and leadership. In: *AMCIS 2016 Proceedings* (2016)
11. Nuottila, J., Aaltonen, K., Kujala, J.: Challenges of adopting agile methods in a public organization. *Int. J. Inf. Syst. Project Manage.* **4**, 65–85 (2016). <https://doi.org/10.12821/ijispm040304>
12. Dikert, K., Paasivaara, M., Lassenius, C.: Challenges and success factors for large-scale agile transformations: a systematic literature review. *J. Syst. Softw.* **119**, 87–108 (2016). <https://doi.org/10.1016/j.jss.2016.06.013>
13. Gregory, P., Barroca, L., Sharp, H., Deshpande, A., Taylor, K.: The challenges that challenge: engaging with agile practitioners' concerns. *Inf. Softw. Technol.* **77**, 92–104 (2016). <https://doi.org/10.1016/j.infsof.2016.04.006>
14. Reunamäki, R., Fey, C.F.: Remote agile: problems, solutions, and pitfalls to avoid. *Bus. Horiz.* **66**, 505–516 (2023). <https://doi.org/10.1016/j.bushor.2022.10.003>
15. Paasivaara, M., Behm, B., Lassenius, C., Hallikainen, M.: Large-scale agile transformation at Ericsson: a case study. *Empir. Softw. Eng.* **23**, 2550–2596 (2018). <https://doi.org/10.1007/s10664-017-9555-8>
16. Sithambaram, J., Nasir, M.H.N.B.M., Ahmad, R.: Issues and challenges impacting the successful management of agile-hybrid projects: a grounded theory approach. *Int. J. Proj. Manag.* **39**, 474–495 (2021). <https://doi.org/10.1016/j.ijproman.2021.03.002>
17. Shameem, M., Kumar, R.R., Kumar, C., Chandra, B., Khan, A.A.: Prioritizing challenges of agile process in distributed software development environment using analytic hierarchy process. *J. Softw. Evol. Process.* **30**, e1979 (2018). <https://doi.org/10.1002/smr.1979>

18. Kamal, T., Zhang, Q., Akbar, M.A.: Toward successful agile requirements change management process in global software development: a client–vendor analysis. *IET Softw.* **14**, 265–274 (2020). <https://doi.org/10.1049/iet-sen.2019.0128>
19. Drivers of Project Agility | PMI. <https://www.pmi.org/learning/thought-leadership/pulse/drive-project-agility>. Accessed 25 Oct 2023
20. Guinan, P.J., Parise, S., Langowitz, N.: Creating an innovative digital project team: levers to enable digital transformation. *Bus. Horiz.* **62**, 717–727 (2019). <https://doi.org/10.1016/j.bushor.2019.07.005>
21. Barke, H., Prechelt, L.: Role clarity deficiencies can wreck agile teams. *PeerJ Comput. Sci.* **5**, e241 (2019). <https://doi.org/10.7717/peerj-cs.241>
22. Fu, M., Tantithamthavorn, C.: GPT2SP: a transformer-based agile story point estimation approach. *IEEE Trans. Software Eng.* **49**, 611–625 (2023). <https://doi.org/10.1109/TSE.2022.3158252>
23. Agile: The Human Factors as the Weakest Link in the Chain | SpringerLink. https://link.springer.com/chapter/10.1007/978-3-319-27896-4_6. Accessed 25 Oct 2023
24. vomBrocke, Jan, Hevner, Alan, Maedche, Alexander: Introduction to design science research. In: vomBrocke, J., Hevner, A., Maedche, A. (eds.) *Design Science Research. Cases*. PI, pp. 1–13. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-46781-4_1
25. Koppenhagen, N., Gaß, O., Müller, B.: *Design Science Research in Action - Anatomy of Success Critical Activities for Rigor and Relevance* (2012). <https://doi.org/10.5445/IR/1000055012>
26. White, J., et al.: A prompt pattern catalog to enhance prompt engineering with ChatGPT. <http://arxiv.org/abs/2302.11382> (2023). <https://doi.org/10.48550/arXiv.2302.11382>
27. Mastering ChatGPT: How to Craft Effective Prompts (Full Guide + Examples). <https://gptbot.io/master-chatgpt-prompting-techniques-guide/>. Accessed 25 Oct 2023
28. DSH, T.: *Mastering Generative AI and Prompt Engineering: A Practical Guide for Data Scientists* [eBook]. <https://datasciencehorizons.com/mastering-generative-ai-prompt-engineering-ebook/>. Accessed 25 Oct 2023
29. OpenAI: GPT-4 Technical Report. <http://arxiv.org/abs/2303.08774> (2023). <https://doi.org/10.48550/arXiv.2303.08774>
30. OpenAI Platform. <https://platform.openai.com/tokenizer>. Accessed 07 June 2023
31. Khan, S.: ChatGPT Memory: Adaptive Prompt Creation. <https://redis.com/blog/chatgpt-memory-project/>. Accessed 25 Oct 2023

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Startup Creation Beyond Hackathons – A Survey on Startup Development and Support

Maria Angelica Medina Angarita¹(✉), Martin Kolnes¹, and Alexander Nolte^{1,2}

¹ University of Tartu, Ülikooli 18, 50090 Tartu, Estonia

{maria.medina, martin.kolnes, alexander.nolte}@ut.ee

² Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Abstract. Hackathons are themed, fast-paced events where participants gather in teams to work on a project of their interest. Hackathons are often organized to drive entrepreneurial behavior, however, little is known about how they have supported startup creation. To address this issue, we conducted a cross-sectional survey among hackathon participants about their motivations for participating in a hackathon including creating a new startup product and advancing their careers. The survey also addressed their perceived hackathon benefits related to entrepreneurship, such as learning and networking, and how useful they were to their startups. Moreover, the survey included aspects of the hackathon setting that may have influenced startup creation, including winning awards. We obtained answers from participants who have attended 48-h, in-person hackathons. We found motivations related to entrepreneurship that were related to startup creation, such as learning about the startup domain. Our findings show that participants with entrepreneurial motivations are more likely to create a startup after the hackathon. We also found that participants with startups in an early stage have attended hackathons motivated to build the initial version of their startup product, however, they have also worked on other projects unrelated to their startup. To support startup creation beyond hackathons, organizers should gain awareness of such hackathon participants who are motivated by entrepreneurship.

Keywords: Entrepreneurial process · Startups · Hackathons

1 Introduction

Hackathons are time-bounded, themed events where participants gather in teams and engage in rapid product development [15, 34]. One area in which hackathons have gained popularity is entrepreneurship. During entrepreneurial hackathons¹, teams are provided with resources including mentorship and awards to encourage them to create startups from their projects [8]. During their early stage of development, startups are newly formed companies faced with immediate challenges regarding establishing a team

¹ We will continue to refer to entrepreneurial hackathons as hackathons.

[20], funding [21], product development [6, 10], and lack of resources [41]. To address these challenges, startup founders have attended incubators, contests, and hackathons [26] as an expression of entrepreneurial behavior. We understand entrepreneurial behavior as a collection of characteristics linked to new venture formation [3]. Prior work in the context of entrepreneurial behavior at hackathons has mainly focused on case studies of individual events which limits the possibility of developing an understanding of how participant motivations can affect startup creation beyond specific contexts [7, 37]. Moreover, preliminary results [30] indicate that some startup founders have attended hackathons after the foundation of their startups. Thus, founders may be motivated to attend hackathons based on the stage of development of their startup [27]. Conversely, participants may not want to create a new startup or develop an existing startup further at the hackathon and attend, instead, for reasons unrelated to startups, such as having fun [24] and free pizza [4]. Thus, we propose our first research question: **RQ1:** How are the motivations of hackathon participants connected to startups?

Developing the hackathon project into a startup project after the hackathon has ended is a main topic of interest in previous research [8]. However, little is known about other entrepreneurial benefits participants have perceived apart from creating a startup at the hackathon, particularly for those participants who already have startups. These benefits include developing the skills of an already existing startup team and getting feedback on an idea related to the startup [25]. We take a broader approach by addressing whether participants were able to create startups after the hackathon ended, and if startup founders with existing startups have brought their startup projects to work on them during the hackathon. Thus, we propose our second research question: **RQ2:** How are the perceived benefits of hackathon participants connected to startups?

Our findings contribute to existing knowledge about the relationship between hackathons and startups by expanding on the motivations and perceived benefits of participants that are related to entrepreneurial behavior and what hackathon aspects may influence startup creation after the hackathon ends.

2 Background

We base our work on findings from two fields: startup research and hackathon research. From the startup research field, we draw on the model of four stages of startup development [20] as it addresses previous frameworks and assigns inherent goals, challenges, and practices to each stage. During the first stage, the *inception* stage, the main goal for founders is to assemble a team to develop a startup product. After the startup product has entered the market, the *stabilization* stage begins, where customer input helps drive the product further. In the next stage, *growth*, the focus switches from product development to business growth, where the main aim is to achieve a significant market share to culminate in *maturity* [20]. Our work contributes to the understanding of how founders of startups in various stages perceive hackathons and their benefits by examining how the motivations (RQ1) and perceived benefits (RQ2) of hackathon participants are connected to startups.

From the hackathon research field, we refer to the motivations (RQ1) and perceived benefits (RQ2) of hackathon participants. Previous research has found that two common motivations (RQ1) are learning and networking [4]. Additional motivations include working with friends who participate [7] and having fun [17, 35]. Little is known about the hackathon motivations of participants that are related to startups. Few studies indicate that they include learning and networking concerning an existing startup, advancing the skills of an already existing startup team [25], and creating a new startup [7, 24]. Our work expands on how these motivations may be connected to a certain startup stage of development. Common hackathon perceived benefits (RQ2) include learning [1, 12], creating technical artifacts [40], and winning awards [7]. In addition, those perceived benefits connected to startups include creating startups [33], learning and networking concerning the startup, and developing the skills of the startup team [25]. Our work contributes to the field of hackathon research by focusing on further perceived benefits related to startups.

2.1 Hypotheses

We propose eight hypotheses (H1–H8) based on our research questions regarding hackathon motivations (RQ1) and perceived benefits for hackathon participants (RQ2).

Hackathon participants commonly focus on developing a product that could become a startup after the hackathon ends [19], therefore, we expect that the most common participant motivations (RQ1) will be related to startup product development (H1). As the main challenge for startups during their *inception* is to build the first version of the product [10, 14, 20, 43], founders with startups at the *inception* stage may be motivated to attend a hackathon to build their startup product if they do not have one (H2). After the period of *stabilization*, when *growth* begins, the main challenge for startups is to achieve a desired growth rate [20], for which there is a need to acquire specialized knowledge and feedback. Thus, founders with startups at later stages may be motivated to attend a hackathon to acquire specialized knowledge and feedback to support their startups (H3).

In addition to the motivations, the creation of startups could be influenced by aspects of the hackathon setting. The quality of the projects developed at the hackathon has been influenced by team size [8], the connection with the stakeholders [13, 22, 32] and the hackathon duration [7, 44]. Learning and productivity have also been found to be influenced by duration [29]. Based on these findings from previous research, we propose that the duration will influence the creation of startups at hackathons (H4).

Prior work about hackathon perceived benefits (RQ2) indicated that founders often built the initial version of their startup product at hackathons [33]. Thus, we propose that founders with startups at the *inception* stage who do not have a startup product will develop it with their team at a hackathon (H5). Moreover, founders who have a startup product have attended a hackathon to learn about topics related to their startups [25]. Thus, we propose that entrepreneurs with startups in later stages will learn about topics related to their startup at a hackathon (H6). However, we do not expect that most hackathon participants have created a startup after a hackathon (H7), as there is little indication of startups being funded after hackathons [30]. Nevertheless, founders may find hackathons the most useful for their startups for product development (H8), as developing an idea into a product in teams is the focus of hackathons.

3 Research Method

The purpose of this study is to identify the motivations of participants to attend hackathons (RQ1), and their perceived benefits (RQ2) to support startup creation at hackathons. As our research method, we used a cross-sectional survey². We selected a survey as our research instrument as it allows for establishing connections and creating a broader overview beyond single events [11]. The survey consisted of various sections that addressed distinct aspects of the research questions (See Table 1). We collected information related to hackathon motivations, and how participants addressed aspects of the hackathon setting in our survey (H1–H4). Considering that some survey participants may have also been startup founders, we asked them if they had founded a startup before or after the hackathon and showed them questions related to their startups in a separate section (H5–H8). Finally, we asked for demographic information such as the age and gender of the participants.

Table 1. Overview of the main survey questions

Aspect	Example item	Based on
Hackathon motivations	“ <i>Creating a new startup</i> ” (Anchored between “ <i>Not at all</i> ” and “ <i>Completely</i> ”)	[11, 25]
Setting (Duration)	(Open-ended)	[2, 8, 25]
Setting (Location)	“ <i>A physically hosted hackathon</i> ” (Single choice)	
Setting (Awards)	“ <i>Opportunity to pitch to investors</i> ” (Multiple choice)	
Project development	“ <i>We analysed the problem we wanted to solve and defined the features to develop</i> ” (Anchored between “ <i>Strongly disagree</i> ” and “ <i>Strongly agree</i> ”. The scales below follow the same format.)	[39]
Learning outcomes	“ <i>I learned about product development</i> ”	[25]
Project satisfaction	“ <i>My ideal outcome towards the hackathon was achieved</i> ”	[36]
Hackathon satisfaction	“ <i>My ideal outcome coming into my project achieved</i> ”	[11]
Hackathon idea	“ <i>Did you bring a startup idea to the hackathon?</i> ” (Single choice)	[25]
Hackathon project	“ <i>Yes, I worked on my startup project</i> ” (Single choice)	[25]

(continued)

² <https://t.ly/dSLn>.

Table 1. (continued)

Aspect	Example item	Based on
Startup team	<i>“Yes, all the members of my startup team were at the hackathon”</i> (Single choice)	[25]
Startup domain	<i>“Software as a service (Saas)”</i> (Single choice)	[38]
Startup stages	<i>“The idea for the startup project was developed but a product had not yet been developed”</i> (Single choice)	[20]
Hackathon usefulness to the startup	<i>“The hackathon was useful to create a product for my startup”</i>	[25]

For our survey, we invited 6142 participants of various 48-h hackathons from 2015 to 2019 in Eastern Europe organized by the same institution. In those hackathons, there was a kickoff at the beginning where participants pitched their ideas and gathered in teams based on the ideas for projects that interested them. They would subsequently work on their projects together while receiving feedback from mentors. In the end, they presented the products they developed at the hackathon, and some teams were awarded prizes, such as funding, to encourage them to continue working on their projects. We obtained 438 responses from the main variables that we submitted to data cleaning. The low number of responses reflects findings from previous research stating that often most survey invites are ignored [5].

3.1 Data Analysis

We carried out a descriptive analysis to gain an understanding of the dataset. This analysis allowed us to determine if founders with startups at the inception stage that did not have a startup product developed it at a hackathon (H5) and whether most participants had created a startup after the hackathon or not (H7). We also created box plots to illustrate the distributions of the variables, such as the perceived hackathon usefulness to the startup (H8). We conducted an exploratory factor analysis using the hackathon motivations (H1) with the Eigenvalues as a reference for determining the number of factors and tested them for inter-item reliability using Cronbach’s α . We chose this test as it measures internal consistency between items on a scale [42]. We also conducted a Mann-Whitney U-test to identify the motivations of startup founders (H2). We chose this test as it allows to find significant statistical differences between two independent variables [23]. Finally, we conducted a logistic regression to find the aspects of the hackathon setting that may have influenced the creation of a startup after the hackathon ended (H4). We did not obtain answers from founders with startups in the growth and maturity stages. Therefore, it is not possible to confirm H3 or H6.

4 Results

We received 438 survey responses of which 164 addressed the main variables used in the statistical analysis. From those 164 responses, we found that 20 respondents marked the *awards* question inaccurately, 3 respondents did not provide any information about the awards they won, 2 respondents marked they had a startup before the hackathon but did not offer any information about them, and 1 responded did not provide data about their startup project. We removed those incomplete responses from the dataset (138).

For the duration of the hackathons, there was a reported minimum of 4 h and a maximum of 72 h. The difference between the 48-h duration and other durations did not allow us to make further statistical analysis with the duration as an aspect of the setting due to the high skewness (H4). Therefore, we conducted further statistical analysis with responses of 48-h hackathons, also known as three-day hackathons (112). Regarding the hackathon setting, 105 (93.75%) respondents marked they attended a physically hosted hackathon, while other respondents marked they attended a hybrid or online hackathon. To avoid imbalance in the dataset we removed all responses from individuals that did not participate in a collocated hackathon. Regarding the demographic of our study participants, there were 68 (64.76%) males, 29 (27.61%) females, 1 (0.95%) non-binary, and 7 (6.66%) participants who abstained from disclosing their gender. Most participants reported being between the ages of 25 to 34 (51.42%), with fewer participants between the ages of 35 to 44 (22.85%), followed by 18 to 24 (18.09%) and 45 to 54 (7.61%).

4.1 Perceived Hackathon Motivations Related to Startups (RQ1)

In this section, we address the hackathon motivations of participants, the factors constituted by different motivations, and the regression analysis.

Hackathon Motivations. We found that *making something cool/working on an interesting project idea* ($\mu = 4.14$, $SD = 0.88$) and *having fun* ($\mu = 4.12$, $SD = 1.01$) were the two most frequent motivations for participants to attend a hackathon, while the least popular motivations were *working on my startup* ($\mu = 2.06$, $SD = 1.40$) and *learning about the domain of my startup* ($\mu = 2.21$, $SD = 1.38$) (see Fig. 1). Thus, our findings do not confirm H1, which states that the most common participant motivations will be associated with startup product development.

We found potential connections between the hackathon motivations using an exploratory factor analysis with varimax rotation. We first performed a Kaiser-Meyer-Okin test to check the suitability of the data, which resulted in a fitting 0.76 value. Based on Eigenvalues, we found five initial factors. We named the factor “*Entrepreneurial*”, and it is constituted by the motivations of *creating a new startup*, *building the first version of a startup product*, *working on my startup*, *developing the skills of my startup team*, *learning about the domain of my startup* and *getting immediate feedback* (See Table 2). We tested the factor for inter-item reliability using Cronbach’s α and found the value of 0.874 acceptable. The second factor, which we named “*Social*”, is constituted by the motivations of *meeting new people* and *becoming part of a community*. We named the following factor “*Achievement*”, it is constituted by the motivations of *winning awards*,

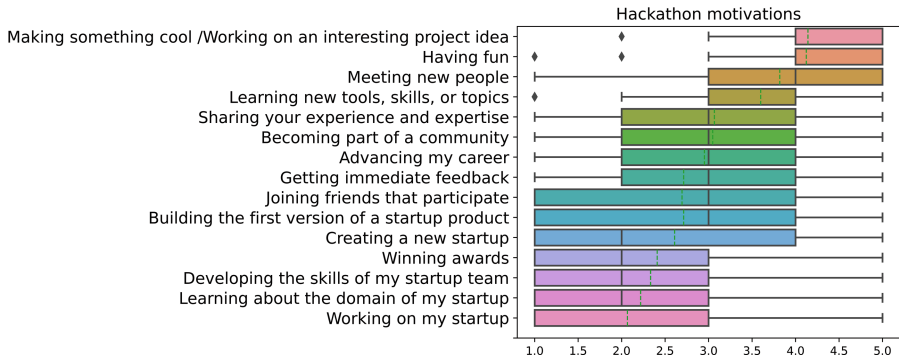


Fig. 1. Motivations of hackathon participants

making something cool/working on an interesting project idea, advancing my career, and sharing your experience and expertise. The following factor is constituted by the motivations of *learning new tools, skills, or topics*, thus, we named it the “*Learning*” factor. Finally, we named the last factor “*Convivial*”, it is constituted by the motivations of *Joining friends that participate* and *Having fun*. We tested these factors and obtained the following Cronbach’s α values: *Social* factor (0.66), *Achievement* factor (0.57), *Learning* factor (n/a), and *Convivial* factor (0.45). As the Cronbach’s α values were insufficient, the remaining factors consist of only one variable: the motivation that scored the highest value for that factor (see highlighted values in Table 2).

Table 2. Exploratory factor analysis. Only values higher than 0.3 for each factor are present.

Motivations and factors	Entrepreneurial	Social	Achievement	Learning	Convivial
Marking something cool/working on an interesting project idea			0.39384		
Learning new tools, skills, or topics				0.94559	
Meeting new people		0.89258			
Sharing your experience and expertise			0.47760		
Advancing my career			0.24662		

(continued)

Table 2. (continued)

Motivations and factors	Entrepreneurial	Social	Achievement	Learning	Convivial
Becoming part of a community		0.52384			
Getting immediate feedback	0.49159				
Joining friends that participate					0.74427
Winning awards			0.67955		
Having fun					0.39587
Creating a new startup	0.76515				
Building the first version of a startup product	0.82695				
Working on my startup	0.7907				
Developing the skills of my startup team	0.65649				
Learning about the domain of my startup	0.67295				

Using a Mann-Whitney U-test, we found that the means of the participants who had founded a startup before or after the hackathon were higher ($\mu = 2.90$) than those who had not ($\mu = 2.67$) for the Entrepreneurial factor ($p < 0.005$). For the founders with a startup at the inception stage without a startup product (14), the Entrepreneurial factor had values of ($\mu = 3.34$, $SD = 0.41$), with the motivation of *building the first version of a startup product* having values of ($\mu = 3.78$, $SD = 1.31$). Thus, confirming H2.

In addition to the motivations, the awards, as an aspect of the hackathon setting, may have influenced startup creation, as they are meant to encourage and support those participants who would like to continue working on their projects. Most of the respondents (74, 70.47%) marked they won an award at the hackathon, while (31, 29.52%) marked they did not. Of the 74 respondents who marked they won an award, some participants reported having won one or more awards: 27 reported they won a team-building experience, 32 indicated that they won a mentoring program, 32 others reported that they won tools and resources, 26 reported they won a cash award, 15 that they won an opportunity to pitch to investors, and 14 reported that they won an award of some other kind.

To identify the motivations or aspects of the hackathon setting that influenced startup creation after the hackathon we conducted a logistic regression (See Table 3). The outcome variable for the regression is *post-hackathon startup formation*, a categorical binary survey item where participants reported yes (1) or no (0) to having founded a startup after the hackathon.

Table 3. Logistic regression results.

Variables	Estimate	SE	OR	p-value
Requirements	0.073	0.423	1.075	0.863
Design	-0.500	0.307	0.607	0.104
Implementation	0.047	0.379	1.048	0.902
Testing	-0.310	0.290	0.734	0.286
Project satisfaction	0.717	0.497	2.048	0.149
Hackathon satisfaction	0.149	0.516	1.161	0.772
Entrepreneurial factor	0.515	0.262	1.674	0.050
Having a startup	0.624	0.795	1.866	0.433
Awards	1.443	0.876	4.232	0.100

Note. The reference category is the response “no” to startup formation. SE = standard error, OR = odds ratio. Requirements to Testing = the degrees of completion of the project

For the predictors, we selected those addressed by previous research about the connection between hackathons and startup formation [25, 31]. They were the awards, the degree of completion of the project (from identifying requirements to testing), the entrepreneurial factor, the perceived hackathon satisfaction, and project satisfaction. We also included having a startup before the hackathon. Along with awards, having a startup is a binary item. The other predictors were survey items that were answered using a five-point Likert scale and later averaged for the regression. The model was statistically significant, $\chi^2(95) = 17.01$, $p = .05$, Cox & Snell [9] $R^2 = 0.15$, Nagelkerke [28] $R^2 = 0.24$ (indicating that 15.0–24.0% of the variance was explained by the model). Sensitivity was 20.0%, and specificity was 98.8%. Out of the nine predictors, one was statistically significant. The entrepreneurial factor predicted startup formation (OR = 1.674, $p = .05$) – a higher entrepreneurial score increased the likelihood of startup formation. However, the confidence in the results is somewhat limited due to the unequal distribution of the dependent variable groups [18] (*startup formation*: 20 = yes; 85 = no). Nevertheless, the results give a preliminary idea about important predictors for startup formation.

4.2 Perceived Hackathon Benefits Related to Startups (RQ2)

In this section, we address the perceived benefits of participants related to startups, the perceived usefulness of the hackathon to the startup, project completion, learning outcomes, satisfaction with the project, and satisfaction with the hackathon.

Of the 105 responses, (92, 87.61%) participants marked they did not have a startup at the time of the hackathon they identified, while only (13, 12.38%) of them did. 29 (27.61%) respondents marked they created a startup before or after the hackathon, among those, 13 marked they created a startup before the hackathon, 20 that they created a startup after the hackathon, and 4 marked they had created a startup before and after the hackathon. Table 4 elaborates on the different startup stages participants reported.

Table 4. Reported startup stages of participants at the time of the hackathon

Startup stages of development	Participants
Inception stage: Startup idea without a startup product	14
Inception stage: With a startup product	4
Stabilization stage: Startup product on the market	1
Other stage	1
No startup idea at the time of the hackathon	9

Most respondents (63, 60%) reported they did not bring a startup idea to the hackathon, while (42, 40%) of them did. Of those 63 participants who did not bring a startup idea to the hackathon, 11 marked they created a startup after the hackathon ended. Of the 42 participants who brought a startup idea to the hackathon, 9 marked they created a startup after the hackathon ended. Only 20 respondents of 105 (19.04%) reported that they created a startup. Thus, supporting H7, as most participants did not create a startup after the hackathon ended. Of the participants that had created a startup before or after the hackathon they attended (29, 27.61%), 12 marked they worked on their startup project after the hackathon, 10 marked they worked on a project that was unrelated to their startup, 5 marked they worked on a project of the same domain of their startup, and 2 marked they worked on other projects.

Of the participants who mentioned that their startup was at the inception stage without a developed product (14, 13.33%), 5 mentioned that they worked on their startup product, other 5 mentioned they worked on a project of their startup domain, and 4 worked on a project unrelated to their startup. Therefore, there is no evidence that confirms H5, as most founders in the inception stage without a startup product did not work on their startup project at the hackathon.

Of the (29, 27.61%) participants who reported they created a startup before or after the hackathon, the most popular startup domain category was *Software as a service* (10), followed by *Others* (8), a Mobile application (4), a Two-sided marketplace (2), E-commerce (3) and media sites (2). Regarding the startup team members, 12 participants marked that there were members of their team at the hackathon, 9 participants that there were no members of their startup team at the hackathon, and 8 reported that all members of the startup team were at the hackathon.

Perceived Usefulness of the Hackathon to the Startup. For the scale of the perceived usefulness of the hackathon to the startup, we analyzed each item individually. The lowest level of agreement was for the statement that the hackathon was useful to create

a product for the startup, pointing toward learning and networking being more useful to startup founders than developing a product at the hackathon (see Fig. 2), thus, rejecting H8.

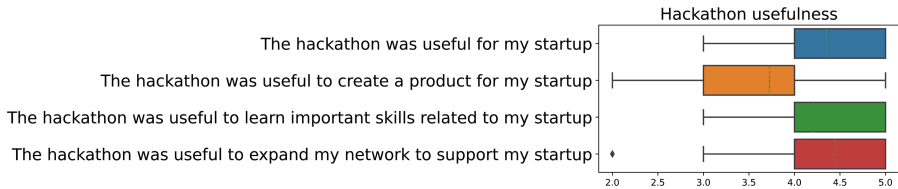


Fig. 2. Perceived hackathon usefulness to the startup

Perceived Project Completion. For this scale, we assigned a description to each of the five stages of the waterfall model (Requirements, design, implementation, verification, and maintenance) [39]. Most participants indicated a high agreement with the first levels of project completion. However, the testing and maintenance processes do not seem to have been conducted as much, with the latter presenting the highest standard deviation (see Fig. 3).

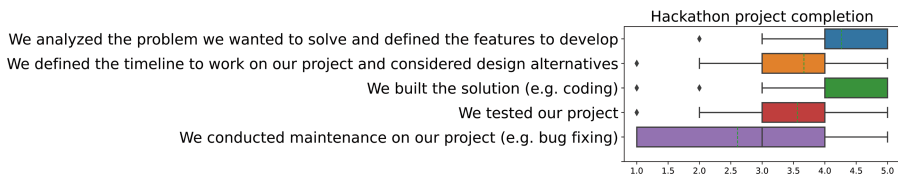


Fig. 3. Perceived degree of project completion

Perceived Hackathon Learning Outcomes. Most participants reported that they learned about product development ($\mu = 3.94$, $SD = 0.93$) and pitching ($\mu = 3.85$, $SD = 1.10$), while the lowest levels of agreement were for learning about the startup domain ($\mu = 3.12$, $SD = 1.20$) and learning how to monetize a product ($\mu = 2.81$, $SD = 1.16$).

Perceived Satisfaction with the Hackathon, and the Project. We tested the scales for *perceived satisfaction with the project* and the *hackathon* for inter-item reliability using Cronbach's α . We found their levels of (0.86) and (0.87) respectively, acceptable to continue to analyze them as one item. Participants indicated an agreement with their perceived satisfaction with the project ($\mu = 3.79$, $SD = 0.88$) and a higher agreement with their perceived hackathon satisfaction ($\mu = 4.12$, $SD = 0.85$).

5 Discussion

We aimed to determine the motivations (RQ1) and perceived benefits (RQ2) of hackathon participants that are related to startups. Table 5 provides an overview of our findings on this relation, including the supported (H2, H7), non-supported (H1, H5, H8), and undetermined (H3, H4, H6) hypotheses.

Table 5. Summary of the hypotheses

Hypotheses	Results
The most common participant motivations will be related to startup product development (H1)	Not supported
Founders with startups at the <i>inception</i> stage may be motivated to attend a hackathon to build their startup product if they do not have one (H2)	Supported
Founders with startups at later stages may be motivated to attend a hackathon to acquire specialized knowledge and feedback to support their startups (H3)	Undetermined
The hackathon duration will influence the creation of startups at hackathons (H4)	Undetermined
Founders with startups at the <i>inception</i> stage that do not have a startup product will develop it at a hackathon (H5)	Not supported
Entrepreneurs with startups in later stages will learn about topics related to their startup at a hackathon (H6)	Undetermined
Most hackathon participants have not created a startup after a hackathon (H7)	Supported
Founders may find hackathons the most useful for their startups for product development (H8)	Not supported

We elaborate on our results from two fields: hackathon research and startup research. Regarding hackathon research, we found that about half of our study participants brought a startup idea to the hackathon, but only a few founded a startup afterward (H7). These findings match those of previous research that reports on challenges that participants face when creating a startup after the hackathon [8, 17]. Thus, it is necessary for hackathon organizers to be aware of those participants who bring startup ideas to the hackathon and to provide them with guidance on what can be done to support their startups after the hackathon ends. We did not obtain answers from founders with startups in later stages (H3, H6). This may suggest that if a founder has a team and a startup product, they may not be interested in engaging in a new project or taking their existing project to a hackathon. Further research may focus on those hackathon aspects that could be useful to founders with startups at later stages.

We also found that the most frequent hackathon motivations (RQ1) are not directly associated with startup product development (H1). The most popular hackathon motivations were, instead, *making something cool/working on an interesting project idea* (*achievement* factor) and *having fun* (*convivial* factor). These findings partially match

previous research where *having fun* [17] was found to be a frequent hackathon motivation. We did, however, find motivations related to entrepreneurship that constituted the *entrepreneurial factor* and reflected diverse aspects of startup development, such as “*Developing the skills of my startup team*” and “*Learning about the domain of my startup*”. Thus, it may seem that participants motivated to create a startup at hackathons are looking forward to addressing multiple challenges of their startup. The *entrepreneurial factor* was also a predictor for startup creation (H4). This finding matches with those from previous research that states that entrepreneurial intention may drive entrepreneurial behavior [16, 19]. Future research about entrepreneurial intention may focus on how to help entrepreneurs stay motivated during the different startup stages and what aspects or challenges of their entrepreneurial journey have demotivated them.

Regarding hackathon perceived learning outcomes (RQ2), we found that participants indicated high levels of learning for pitching and product development, but less so for learning how to monetize a product, and the domain of their startup. These findings match those of previous research where pitching was reported amongst the most popular topics addressed at the hackathon [25] and where participants learned within their teams “*from doing*” in situ [12].

Regarding the startup research field, we found that although some startup founders have attended hackathons motivated to work on the first version of their startup product (H2), and some have developed their startup products, or projects related to its domain (H5), the least perceived usefulness to the startup was in creating the startup product at the hackathon (H8). This finding points toward participants not perceiving the project developed at the hackathon to be necessarily suitable for their startup.

Previous research has also pointed toward participants not developing their startup product at the hackathon [25]. This finding may be related to the fact that our study participants reported low levels of agreement with the testing and maintenance of their projects (RQ2). They may not be motivated to use the hackathon project as their startup project, as it may lack maturity. Conversely, the reported low levels of agreement with the testing and maintenance of the projects may also be related to the duration [44] or the lack of previously developed projects at the hackathon. Valuing other benefits over the development of a project is also supported by the high level of agreement with the satisfaction with the hackathon compared to the satisfaction with the project (RQ2).

5.1 Limitations

Our research was based on an online survey that addressed the individual experiences of hackathon participants with a focus on their perceptions and opinions. However, certain aspects of the hackathon setting that may have influenced the perceived benefits were unobserved. For the process of working in teams, such aspects include goal clarity, the match between skills and tasks, and satisfaction with the team process. We could not observe these aspects as the study participants attended different hackathons, thus we focused on individual perceptions instead. Moreover, it is unknown if the 105 survey participants are a representative cross-section of the overall hackathon population, as we studied events in a specific geographic context organized by the same institution. We accepted this limitation because studying similar events allowed us to assume similar

settings in which they were obtained. Our findings are limited to the setting and participants we studied and future research in a different context may yield different results. We also created questionnaire items ourselves that may pose a threat to reliability and validity, we did, however, not use them for any statistical analysis as combined scales.

6 Conclusion

Our findings suggest that many hackathon participants brought a startup idea to a hackathon, and some of them also had motivations related to startup creation that are part of the *entrepreneurial factor*, a predictor for startup creation. Thus, startup creation can be supported at hackathons when organizers are aware of the entrepreneurial motivations of the participants [24]. This awareness can begin when participants report to the organizers their motivations as they register for the hackathon. The motivation of participants could potentially influence how they work together in teams, as teams where participants have different motivations could have more difficulties aligning their goals. During the planning of a hackathon, organizers should consider the motivations and needs that the participants express, including those apart from collaborative product development, such as learning and networking.

Acknowledgment. The first author's contribution to this project is partially funded by the Creative Impact Research Centre Europe (CIRCE).

References

1. Affia, A.O., Nolte, A., Matulevičius, R.: Developing and evaluating a hackathon approach to foster cyber security learning. In: Nolte, A., Alvarez, C., Hishiyama, R., Chounta, I.-A., Rodríguez-Triana, M.J., Inoue, T. (eds.) CollabTech. LNCS, vol. 12324, pp. 3–19. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58157-2_1
2. Akematsu, Y., Yoshino, T.: Proposal of emphasized pseudo expression for improving the recognition of the presence and contribution of remote participants in cooperative work. In: Nakanishi, H., Egi, H., Chounta, I.-A., Takada, H., Ichimura, S., Hoppe, U. (eds.) CRIWG+CollabTech. LNCS, vol. 11677, pp. 262–271. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28011-6_20
3. Bezerra, C.M.D.S. et al.: Entrepreneurial behavior and strategy: a systematic literature review. REGEPE Entrep. Small Bus. J., e2396 (2023). <https://doi.org/10.14211/regepe.esbj.e2396>
4. Briscoe, G., Mulligan, C.: Digital Innovation: The Hackathon Phenomenon, p. 13. Creative-works, London (2014)
5. Chounta, I.-A., Nolte, A.: The CAT effect: exploring the impact of casual affective triggers on online surveys' response rates. In: CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2022). <https://doi.org/10.1145/3491102.3517481>
6. Cico, O., Souza, R., Jaccheri, L., Nguyen Duc, A., Machado, I.: Startups transitioning from early to growth phase - a pilot study of technical debt perception. In: Klotins, E., Wnuk, K. (eds.) ICSOB 2020. LNBIP, vol. 407, pp. 102–117. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67292-8_8

7. Cobham, D., et al.: From Appfest to entrepreneurs: using a hackathon event to seed a university student-led enterprise. Presented at the International Technology, Education and Development Conference, Valencia, Spain, March 2017
8. Cobham, D., et al.: From hackathon to student enterprise: an evaluation of creating successful and sustainable student entrepreneurial activity initiated by a university hackathon. Presented at the International Conference on Education and New Learning Technologies, Barcelona, Spain, March 2017
9. Cox, D.R., et al.: Analysis of Binary Data. Chapman & Hall [u.a.], Boca Raton (1999)
10. Crowne, M.: Why software product startups fail and what to do about it. Evolution of software product development in startup companies. In: IEEE International Engineering Management Conference, Cambridge, pp. 338–343. IEEE (2002)
11. Filippova, A., et al.: From diversity by numbers to diversity as process: supporting inclusiveness in software development teams with brainstorming. In: 2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE), Buenos Aires, pp. 152–163. IEEE (2017). <https://doi.org/10.1109/ICSE.2017.22>
12. Gama, K., et al.: Hackathons in the formal learning process. In: Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018, Larnaca, Cyprus, pp. 248–253. ACM Press (2018)
13. Gama, K., et al.: Mapathons and hackathons to crowdsource the generation and usage of geographic data. In: Proceedings of the International Conference on Game Jams, Hackathons and Game Creation Events 2019 - ICGJ 2019, San Francisco, CA, USA, pp. 1–5. ACM Press (2019)
14. Giardino, C., Wang, X., Abrahamsson, P.: Why early-stage software startups fail: a behavioral framework. In: Lassenius, C., Smolander, K. (eds.) ICSOB 2014. LNBP, vol. 182, pp. 27–41. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08738-2_3
15. Happonen, A., et al.: Study on hackathons for new innovation seed and business model development needs in digitalization driven sustainability, circularity and environmentally friendly solutions demanding digitalizing societies. In: Kuo, Dr.C.-C. (ed.) New Innovations in Economics, Business and Management, vol. 4, pp. 1–29. Book Publisher International (a part of SCIENCEDOMAIN International) (2022)
16. Joensuu-Salo, S., et al.: Do intentions ever die? The temporal stability of entrepreneurial intention and link to behavior. ET **62**(3), 325–338 (2020)
17. Juell-Skielse, G., Hjalmarsson, A., Johannesson, P., Rudmark, D.: Is the public motivated to engage in open data innovation? In: Janssen, M., Scholl, H.J., Wimmer, M.A., Bannister, F. (eds.) EGOV 2014. LNCS, vol. 8653, pp. 277–288. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44426-9_23
18. King, G., Zeng, L.: Logistic regression in rare events data. Polit. Anal. **9**(2), 137–163 (2001)
19. Kitsios, F., Kamariotou, M.: Open data hackathons: an innovative strategy to enhance entrepreneurial intention. IJIS **10**(4), 519–538 (2018)
20. Klotins, E., et al.: A progression model of software engineering goals, challenges, and practices in start-ups. IEEE Trans. Software Eng. **47**(3), 498–521 (2021)
21. Konsek-Ciechońska, J.: Startup companies, challenges in Poland. KIJ **30**(6), 1621–1626 (2019). <https://doi.org/10.35120/kij30061621K>
22. Linnell, N., et al.: Hack for the homeless: a humanitarian technology hackathon. In: IEEE Global Humanitarian Technology Conference (GHTC 2014), San Jose, CA, pp. 577–584. IEEE (2014). <https://doi.org/10.1109/GHTC.2014.6970341>
23. Luomakoski, J., et al.: Innovation hackathon what did the students get from it? Presented at the Entrepreneurship Education Conference, Finland (2021)
24. Medina Angarita, M.A., Nolte, A.: Does it matter why we hack? – Exploring the impact of goal alignment in hackathons. In: 17th European Conference on Computer-Supported Cooperative Work, vol. 16 (2019)

25. Medina Angarita, M.A., Nolte, A.: Supporting entrepreneurship with hackathons - a study on startup founders attending hackathons. In: Wang, X., Martini, A., Nguyen-Duc, A., Stray, V. (eds.) ICSOB 2021. LNBIP, vol. 434, pp. 107–121. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-91983-2_9
26. Merguei, N., Costa, C.: What are pre-acceleration programs? *J. Bus. Ventur. Insights* **18**, e00324 (2022)
27. Miranda, A., et al.: An exploratory study of the benefits of time-bounded collaborative events for startup founders. In: 2023 IEEE/ACM 16th International Conference on Cooperative and Human Aspects of Software Engineering (CHASE), Melbourne, Australia, pp. 43–55. IEEE (2023)
28. Nagelkerke, N.J.D.: Miscellanea A note on a general definition of the coefficient of determination. *Biometrika* **78**(3), 691–692 (1991)
29. Nandi, A., Mandernach, M.: Hackathons as an informal learning platform. In: Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE 2016, Memphis, Tennessee, USA, pp. 346–351. ACM Press (2016)
30. Nolte, A.: Touched by the hackathon: a study on the connection between hackathon participants and start-up founders. In: Proceedings of the 2nd ACM SIGSOFT International Workshop on Software-Intensive Business: Start-ups, Platforms, and Ecosystems - IWSiB 2019, Tallinn, Estonia, pp. 31–36. ACM Press (2019)
31. Nolte, A., et al.: What happens to all these hackathon projects? – Identifying factors to promote hackathon project continuation. *Proc. ACM Hum.-Comput. Interact.* **4**, 26 (2020)
32. Nolte, A., et al.: You hacked and now what?: - exploring outcomes of a corporate hackathon. *Proc. ACM Hum.-Comput. Interact.* **2**(CSCW), 1–23 (2018)
33. Olson, K.R., et al.: Health hackathons: theatre or substance? A survey assessment of outcomes from healthcare-focused hackathons in three countries. *BMJ Innov.* **3**(1), 37–44 (2017)
34. Pe-Tham, E.P.P., et al.: Designing corporate hackathons with a purpose: the future of software development. *IEEE Softw.* **36**(1), 15–22 (2019)
35. Purwanto, A., et al.: Citizens’ motivations for engaging in open data hackathons. In: Panagiotopoulos, P., et al. (eds.) *Electronic Participation*, pp. 130–141. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-27397-2_11
36. Qiu, H.S. et al.: Approaches to diversifying the programmer community – the case of the girls coding day. In: 2021 IEEE/ACM 13th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), Madrid, Spain, pp. 91–100. IEEE (2021)
37. Richter, N., Dragoeva, D.: Digital entrepreneurship and agile methods—a hackathon case study. In: Soltanifar, M., et al. (eds.) *Digital Entrepreneurship: Impact on Business and Society*, pp. 51–68. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-53914-6_3
38. Rompho, N.: Operational performance measures for startups. *MBE* **22**(1), 31–41 (2018)
39. Ruparelia, N.B.: Software development lifecycle models. *SIGSOFT Softw. Eng. Notes.* **35**(3), 8–13 (2010)
40. Soltani, P., et al.: Hackathon – a method for digital innovative success: a comparative descriptive study. In: Proceedings of the 8th European Conference on Information Management and Evaluation, ECIME 2014 (2014)
41. Sutton, S.M.: The role of process in software start-up. *IEEE Softw.* **17**(4), 33–39 (2000)
42. Tavakol, M., Dennick, R.: Making sense of Cronbach’s alpha. *Int. J. Medical Educ.* **2**, 53–55 (2011). <https://doi.org/10.5116/ijme.4dfb.8dfd>
43. Wang, X., Edison, H., Bajwa, S.S., Giardino, C., Abrahamsson, P.: Key challenges in software startups across life cycle stages. In: Sharp, H., Hall, T. (eds.) *XP 2016*. LNBIP, vol. 251, pp. 169–182. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-33515-5_14
44. Wilson, J., et al.: Beyond the classroom: the impact of a university-based civic hackathon addressing homelessness. *J. Soc. Work. Educ.* **55**(4), 736–749 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Starting Collaborations Between SMEs and Researchers in Software Engineering

Sergio Rico^(✉), Felix Dobsław, and Lena-Maria Öberg

Department of Communication, Quality Management, and Information Systems,
Mid Sweden University, östersund, Sweden
{sergio.rico,felix.dobslaw,lena-maria.oberg}@miun.se

Abstract. In software engineering research, academia-industry collaboration is predominantly understood as partnerships between academic institutions and large companies. Small and Medium-sized Enterprises (SMEs) are vital contributors to the industry, and they are numerous. Their unique preconditions and challenges differentiate their collaboration dynamics from larger corporations. We seek to identify guiding principles and practices for initiating collaborations between researchers and SMEs. Through a meta-synthesis approach drawn from two systematic literature reviews, we introduce a collaborative model canvas. This emphasizes the importance of SMEs' business contexts and the relationships between researchers and SMEs. Our research offers insights for those looking to collaborate with SMEs, considering potential challenges and limitations.

Keywords: industry collaboration · SMEs · software engineering

1 Introduction

Industry-academia collaboration in software engineering is fundamental for successful research, fostering win-win relationships [4]. These collaborations grant academic researchers access to real-world problems and data for empirical validation and align with universities' mission to drive regional economic and social development [9]. Moreover, such a hands-on approach enhances academic programs with practical insights [29]. For businesses, this collaboration connects research outcomes tailored to their challenges, facilitates upskilling and reskilling, and provides a gateway to recruit students [5]. Collaboration can push regional development and economic growth [2, 9].

Research on industry-academia collaboration in software engineering has mainly been centered around large companies [12, 16, 27, 35], with the collaboration involving small and medium-sized enterprises (SMEs) receiving considerably less attention. Particularly in northern Nordic regions such as Finland, Norway, and Sweden, SMEs form a substantial part of the software landscape, with a pronounced tilt towards consulting and services rather than in-house development [30]. Unlike their larger counterparts, SMEs face challenges like limited resources [23] and cognitive barriers [8]. With the rapid pace of digitalization and AI advancements, the pressure on SMEs to stay at the forefront is

high. In this rapidly changing landscape, institutions like ours, providing software engineering and information systems programs, recognize the importance of collaborating with regional SMEs. Engaging in these partnerships confirms our academic endeavors align with these enterprises' real-world challenges.

Our study reinterprets existing literature to address the practical challenges of initiating collaborations between researchers and SMEs. Utilizing a qualitative meta-synthesis approach [18], we delve into two notable Systematic Literature Reviews (SLRs) [2, 12]. From this analysis, we synthesize a Collaborative Model Canvas as a tool designed to foster collaboration between researchers and SMEs in software engineering. While primarily targeting researchers, the canvas offers insights for SMEs, local governments, and universities, highlighting the challenges and potentials of these collaborative partnerships. The following questions drive our study:

RQ1: What distinguishes collaborations with SMEs from those with large companies, and what challenges are unique to SME collaborations?

RQ2: Which insights from previous research on industry-academia collaborations can be adapted for collaborations between researchers and SMEs?

2 Background and Related Work

SMEs are crucial to the global economy. For instance, 99% of all EU businesses are SMEs, providing two-thirds of private sector jobs [31]. Innovation and research play a vital role in the growth and competitiveness of these SMEs. Research in software engineering has explored best practices for SMEs [1] and examined challenges and best practices of software startups [14]. While software startups focus on scalable software-based products or services, their challenges upon scaling are similar to those encountered by SMEs [20].

Collaborating with SMEs offers unique opportunities compared to larger organizations [23], but it also implies challenges. Within the regional innovation ecosystem, which encompasses SMEs, startups, regional authorities, and third parties like incubators and science parks, several factors influence these collaborations. Specifically, SMEs often face resource limitations, preventing them from engaging in sustained research collaborations [23]. The absence of pre-existing research connections complicates initiating collaborative projects for SMEs, which often lack established networks with research institutions [8]. Moreover, limited exposure to research and innovation may hinder SMEs' recognition of the value of collaborations, affecting strategic planning for partnerships with researchers [6].

Although industry-academia collaboration in software engineering has received attention in the literature [12], most research targets large companies, such as the technology transfer model [13] and the agile collaborative approach [28]. However, some frameworks, including the Certus [16] and Continuous Collaborative [17] models, incorporate SMEs, though not as a central part of the collaboration. Our study contributes to filling this gap by adapting and applying literature-derived insights to the unique context of SMEs.

3 Research Methodology: Meta-Synthesis of SLRs

To address our research questions, we adopted a meta-synthesis approach [18], focusing on an interpretative paradigm. This synthesis sought to derive actionable insights for SMEs using data from two chosen SLRs [2, 12].

1. **Source Selection:** We analyzed two comprehensive SLRs. Ankrah et al.'s SLR provides a holistic view of university-industry collaborations, detailing motivations, challenges, and practices [2]. Conversely, Garousi et al.'s SLR focuses on software engineering collaboration challenges and practices [12].
2. **Interpretative and Comparative Analysis:** Drawing from our experiences with SME collaborations, we extracted and systematically analyzed data from the selected SLRs. Our focus centered on Organizational Forms [2], Motivations [2], Challenges [12] and Best Practices [12].
3. **Synthesis and Model Development:** We designed the Collaborative Model Canvas from our analyses, taking inspiration from the Business Model Canvas [19].
4. **Feedback and Discussion:** After drafting the Collaborative Model Canvas, we shared it online, refining it based on co-author discussions.

Our methodology has certain limitations. It relies on two SLRs that are few years old. To our knowledge, no recent secondary studies have examined either industry-academia collaboration or the role of SMEs, underlining the significance of our research. The broader industry-academia collaborations might not fully cover the unique dynamics of SMEs and startups. Potential biases from our perspectives and experiences underline the need for further empirical validations.

4 Collaborative Model Canvas

The Collaborative Model Canvas, detailed in Fig. 1, is a framework to guide the initiation of collaborations between researchers and SMEs. It outlines crucial considerations for collaboration yet remains adaptable, permitting customization, e.g., based on the expertise area of researchers. This canvas is not prescriptive. Instead, it offers a starting point to design and initiate collaborations.

4.1 Partners

Beyond researchers and SMEs, third parties can be essential in promoting and facilitating collaboration [22]. We identified various stakeholders: universities, local government, incubators, accelerators, technology transfer offices, company associations, and entrepreneurs. While researchers provide academic rigor, SMEs contribute with real-world challenges. Regional governments aim to enhance economic and technological development by fostering closer collaborations between researchers and SMEs [32, 33].







<p>Partners </p> <p><i>Stakeholders involved in the collaboration.</i></p> <p>Academic researchers SMEs Regional government organizations Startups External researchers and research groups Incubators, accelerators, technology transfer offices National initiatives</p>	<p>Value Propositions </p> <p><i>Goals that the collaboration aims to achieve.</i></p> <p>Applied solutions for SMEs based on research outcomes Data and real problems for researchers Network development Win-win benefit New product and MVP development Balance academic rigor with business relevance</p>	<p>Channels and Activities </p> <p><i>Core activities undertaken within the scope of the collaboration.</i></p> <p>Regular networking Joint formulation of problems and research questions Training courses MVP and prototype development Design science Knowledge dissemination and publication Small iterative, incremental projects</p>
<p>Relationships </p> <p><i>Principles and practices that sustain personal and organizational relationships.</i></p> <p>Open and regular communication Mutual respect and understanding Avoid complicated jargon Simple management Long-term relationship Entrepreneurs' engagement. Champions</p>	<p>Benefits </p> <p><i>Tangible and intangible rewards yielded by the collaboration</i></p> <p>SMEs: Improved business processes or products (tools and code); Awareness Researchers: Research publications; Insights for teaching; Grants Business ecosystem: New business ideas Students: More practical courses; higher employability Region: Economic growth</p>	<p>Resources and Costs </p> <p><i>Assets and resources required for the collaboration.</i></p> <p>Funding (Grants, SME investments) Decision-makers in SMEs SME resource allocation Business scenarios, pain points, and needs from SMEs Time and Effort</p>

Fig. 1. Collaborative Model Canvas with key components. See [Supplementary Material](#) for an expanded view and key practices for each component.

Governmental offices and agencies are also potential partners, as the fields of software, digitalization, and AI are increasingly crucial to the operations of government offices and agencies [34]. Incubators and accelerators can play a role when academic researchers are involved in helping to develop or validate new products or services and in the founding of startups [7].

Individuals, especially researchers, play a crucial role in initiating and fostering partnerships between academia and SMEs [2,23]. Entrepreneurs and SME leaders, deeply integrated into daily operations, influence decision-making significantly, making their active engagement essential for successful collaboration [12].

4.2 Value Proposition

The model’s value proposition focuses on achieving mutual benefits through a blend of academic rigor, business relevance, and practicality [10]. Collaborations should prioritize the immediate challenges of SMEs, given their low failure tolerance, while setting the stage for long-term partnerships. Emphasizing short-term gains and sustained collaboration is vital, as it aligns with the SMEs’ immediate needs and drives for adaptation and innovation [23].

4.3 Channels and Activities

The following channels were identified when initiating collaborative initiatives:

- Personal Relationships: Initial touchpoints that foster trust between researchers and SMEs [21].
- Research Projects: Structured settings for deep collaboration with specific focus areas.

- Education and Training: Courses for SMEs, workshops, and informal hackathons promoting training and knowledge sharing. University courses incorporate real-world issues, with SME guest lectures enhancing practicality [11].
- Local Business Ecosystem: Encompasses SMEs, startups, incubators, accelerators, and government entities [22].
- Researchers’ Role in Business: Assisting in validating concepts and prototyping for startups [7].

Activities within the collaborative framework refer to the “what”, or the tasks and actions undertaken. These activities should be conducted iteratively and incrementally to minimize risks and deliver value in both the short and long term [27]. Key activities include co-formulating research questions that align with SME operational challenges, applying for joint research grants, and undertaking practical steps like testing and piloting [12,13]. These activities aim to ensure the collaboration’s financial and practical sustainability and the research outcomes’ applicability. Furthermore, knowledge dissemination offers a chance to encourage dialogue. It involves not only publishing in academic journals but also engaging with wider audiences through blogs, webinars, and social media, enhancing visibility within and outside academic context [3].

Case studies, action research, and design science are methodologies to consider when collaborating with SMEs. Design science, in particular, allows researchers to address similar challenges and design interventions beneficial for similar contexts [25].

4.4 Collaborative Relationships

We have identified five key principles for establishing and maintaining collaborative relationships between researchers and SMEs. First, building and nurturing personal relationships are vital in the collaboration between researchers and SMEs. Beyond the organizational boundaries, personal relationships must be nurtured and maintained to ensure the active participation of all stakeholders and the longevity of the collaboration [26]. Second, the collaboration should aim to develop long-term relationships within the ecosystem [12,26]. The time horizons of SMEs and researchers differ, but the collaboration with SMEs should be envisioned as a long-term relationship. Third, maintaining open and regular communication is key to building trust, aligning with SMEs’ needs, and clarifying the management of intellectual property rights [35]. Fourth, envision the collaboration as a win-win, where both entities benefit mutually [4]. Lastly, the presence of champions within SMEs is essential. Champions are engaged, well-networked, and deeply committed to the project, effectively communicating its benefits to decision-makers [35].

4.5 Benefits

SMEs benefit from tailored solutions resulting in improved business processes or products, often materializing as tools or code [24]. Researchers gain from applied

research opportunities, avenues for publications, and potential funding, thereby adding legitimacy to their academic work [2].

Universities see a dual benefit: the enrichment of educational content and the increased involvement of students in real-world projects. This educational approach enriches the curriculum and enhances students' employability, providing practical experience closely aligned with industry needs [5, 11].

Local economies and employment benefit from these collaborations. They spur innovation and growth and introduce new business ideas, fostering economic advancement and community enhancement. Additionally, SMEs can network with students, facilitating recruitment and access to the latest skill sets [8, 15].

4.6 Resources and Costs

Key resources include funding avenues such as grants, SME investments, and other financial mechanisms like government initiatives [8]. While SMEs might not directly fund research, their participation in grant applications can improve financial viability. Effective resource management is crucial for research activities and real-world implementation, impacting the collaboration's long-term sustainability and success [2, 35].

On the other hand, the collaboration also incurs various costs. Time investments are significant for building relationships, facilitating communication, and organizing events like workshops. Resource expenditures are not solely financial but involve the human and intellectual capital needed to sustain the collaboration and execute incremental projects [2]. Additional costs may emerge, such as those for on-site activities and the continuous alignment of the research focus with SMEs' evolving needs.

5 Conclusion

In addressing **RQ1**, our exploration highlights the distinct dynamics and challenges SMEs face when collaborating with researchers compared to larger companies. SME collaborations often involve more stakeholders, such as regional government bodies, technology transfer offices, and universities. These groups play a crucial role in enabling collaborations, a factor especially critical for SMEs who may be constrained by limited resources and narrower knowledge networks. Research relevance becomes essential for SMEs, who typically prioritize immediate outcomes and might hesitate to commit to extensive research engagements without guaranteed short-term benefits. In the SME setting, the absence of formalized research infrastructures emphasizes the need for robust interpersonal trust and clear communication. While collaborations with large corporations may be more direct, SME partnerships can span a diverse range, from educational initiatives to startup businesses or product validations.

For **RQ2**, our literature examination revealed key insights about industry-academia collaborations adaptable to the SME context. Collaborations arise

from planning, commitment, and researchers' active roles in initiating partnerships. While established frameworks may guide industry-academia collaboration, they need adaptation for SME-specific challenges and opportunities. Maintaining relevant research outcomes and open communication are vital for success. Our work also highlights the value of meta-research in advancing SMEs-researchers collaboration.

This paper explores researchers-SME collaborations in software engineering, drawing from existing literature to outline guiding principles and practices. We introduce the collaborative model canvas as a comprehensive framework to assist researchers and SMEs in starting joint projects. The canvas may serve as a roadmap for researchers and provide SMEs access to research outcomes. There is a need for researchers who lead these collaborations and fostering relationships with SMEs. Additionally, our work highlights the significant benefits of such collaborations, suggesting that educational institutions and governments should invest in them to promote education and boost local economies. Future research should focus on empirically assessing the canvas to facilitate collaborations with SMEs, refine the framework, and investigate potential avenues for industry-academia collaboration with SMEs.

Supplementary Material: <https://doi.org/10.5281/zenodo.10093192>

References

1. Abrahamsson, P., et al.: Software startups-a research agenda. *e-Informatica Softw. Eng. J.* **10**(1), 1–29 (2016)
2. Ankrah, S., Omar, A.T.: Universities-industry collaboration: a systematic review. *Scand. J. Manag.* **31**(3), 387–408 (2015)
3. Baldassarre, M.T., Caivano, D., Visaggio, G.: Empirical studies for innovation dissemination: ten years of experience. In: *Proceedings of the 17th International Conference on Evaluation and Assessment in Software Engineering* (2013)
4. Basili, V., Briand, L., Bianculli, D., Nejati, S., Pastore, F., Sabetzadeh, M.: Software engineering research and industry: a symbiotic relationship to foster impact. *IEEE Softw.* **35**(5), 44–49 (2018)
5. Beckman, K., Coulter, N., Khajenoori, S., Mead, N.R.: Collaborations: closing the industry-academia gap. *IEEE Softw.* **14**(6), 49–57 (1997)
6. Bigliardi, B., Galati, F.: Which factors hinder the adoption of open innovation in SMEs? *Technol. Anal. Strateg. Manage.* **28**(8), 869–885 (2016)
7. Bjarnason, E., Lang, F., Mjöberg, A.: An empirically based model of software prototyping: a mapping study and a multi-case study. *Empir. Softw. Eng.* **28**(5), 115 (2023)
8. Carayannis, E.G., Meissner, D.: Glocal targeted open innovation: challenges, opportunities and implications for theory, policy and practice. *J. Technol. Transf.* **42**(2), 236–252 (2017)
9. Compagnucci, L., Spigarelli, F.: The third mission of the university: a systematic literature review on potentials and constraints. *Technol. Forecast. Soc. Chang.* **161**, 120284 (2020)

10. Garousi, V., Borg, M., Oivo, M.: Practical relevance of software engineering research: synthesizing the community's voice. *Empir. Softw. Eng.* **25**, 1687–1754 (2020)
11. Garousi, V., Giray, G., Tuzun, E., Catal, C., Felderer, M.: Closing the gap between software engineering education and industrial needs. *IEEE Softw.* **37**(2), 68–77 (2019)
12. Garousi, V., Petersen, K., Ozkan, B.: Challenges and best practices in industry-academia collaborations in software engineering: a systematic literature review. *Inf. Softw. Technol.* **79**, 106–127 (2016)
13. Gorschek, T., Garre, P., Larsson, S., Wöhlín, C.: A model for technology transfer in practice. *IEEE Softw.* **23**(6), 88–95 (2006)
14. Larrucea, X., O'Connor, R.V., Colomo-Palacios, R., Laporte, C.Y.: Software process improvement in very small organizations. *IEEE Softw.* **33**(2), 85–89 (2016)
15. Leckel, A., Veilleux, S., Dana, L.P.: Local open innovation: a means for public policy to increase collaboration for innovation in SMEs. *Technol. Forecast. Soc. Chang.* **153**, 119891 (2020)
16. Marijan, D., Gotlieb, A.: Industry-academia research collaboration in software engineering: the Certus model. *Inf. Softw. Technol.* **132**, 106473 (2021)
17. Mikkonen, T., Lassenius, C., Männistö, T., Oivo, M., Järvinen, J.: Continuous and collaborative technology transfer: software engineering research with real-time industry impact. *Inf. Softw. Technol.* **95**, 34–45 (2018)
18. Noah Jr, P.D.: A systematic approach to the qualitative meta-synthesis. *Issues Inf. Syst.* **18**(2) (2017)
19. Osterwalder, A., Pigneur, Y.: *Business Model Generation: a Handbook for Visionaries, Game Changers, and Challengers*, vol. 1. Wiley, Hoboken (2010)
20. Paternoster, N., Giardino, C., Unterkalmsteiner, M., Gorschek, T., Abrahamsson, P.: Software development in startup companies: a systematic mapping study. *Inf. Softw. Technol.* **56**(10), 1200–1218 (2014)
21. Perkmann, M., Walsh, K.: University-industry relationships and open innovation: towards a research agenda. *Int. J. Manag. Rev.* **9**(4), 259–280 (2007)
22. Radziwon, A., Bogers, M.: Open innovation in SMEs: exploring inter-organizational relationships in an ecosystem. *Technol. Forecast. Soc. Chang.* **146**, 573–587 (2019)
23. Rajalo, S., Vadi, M.: Collaboration potential between low-capacity SMEs and academic researchers determined by symmetry of motivation. *Technovation* **107**, 102304 (2021)
24. Rico, S., Bjarnason, E., Engström, E., Höst, M., Runeson, P.: A case study of industry-academia communication in a joint software engineering research project. *J. Softw. Evol. Process* **33**(10), e2372 (2021)
25. Runeson, P., Engström, E., Storey, M.-A.: The Design Science Paradigm as a Frame for Empirical Software Engineering. In: *Contemporary Empirical Methods in Software Engineering*, pp. 127–147. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-32489-6_5
26. Runeson, P., Minör, S.: The 4+ 1 view model of industry-academia collaboration. In: *Proceedings of the 2014 International Workshop on Long-Term Industrial Collaboration on Software Engineering*, pp. 21–24 (2014)
27. Sandberg, A., Pareto, L., Arts, T.: Agile collaborative research: action principles for industry-academia collaboration. *IEEE Softw.* **28**(4), 74–83 (2011)
28. Sandberg, A.B., Crnkovic, I.: Meeting industry-academia research collaboration challenges with agile methodologies. In: *39th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP)* (2017)

29. Serow, R.C.: Research and teaching at a research university. *High. Educ.* **40**, 449–463 (2000)
30. Terås, J.T., Myhr, S., Nygaard, V.: *Socio-economic analysis: Interreg aurora*. Technical Report, NORCE Norwegian Research Centre AS (2021)
31. The European Commission: *Unleashing the full potential of European SMEs* (2020)
32. Tödting, F., Kaufmann, A.: The role of the region for innovation activities of SMEs. *Eur. Urban Reg. Stud.* **8**(3), 203–215 (2001)
33. Tödting, F., Kaufmann, A.: SMEs in regional innovation systems and the role of innovation support-the case of upper Austria. *J. Technol. Transf.* **27**(1), 15–26 (2002)
34. Toll, D., Lindgren, I., Melin, U., Madsen, C.Ø.: Artificial intelligence in Swedish policies: Values, benefits, considerations and risks. In: *Electronic Government: 18th IFIP WG 8.5 International Conference, EGOV 2019* (2019)
35. Wohlin, C., et al.: The success factors powering industry-academia collaboration. *IEEE Softw.* **29**(2), 67–73 (2011)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Towards a Business Case for AI Ethics

Mamia Agbese¹[✉], Erika Halme¹, Rahul Mohanani¹,
and Pekka Abrahamsson²

- ¹ The University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä, Finland
{maoragbe,anermesi,rahul.p.mohanani}@jyu.fi
² Tampere University, Kalevantie 4, 33100 Tampere, Finland
pekka.abrahamsson@tuni.fi

Abstract. The increasing integration of artificial intelligence (AI) into software engineering (SE) highlights the need to prioritize ethical considerations within management practices. This study explores the effective identification, representation, and integration of ethical requirements guided by the principles of IEEE Std 7000–2021. Collaborating with 12 Finnish SE executives on an AI project in autonomous marine transport, we employed an ethical framework to generate 253 ethical user stories (EUS), prioritizing 177 across seven key requirements: traceability, communication, data quality, access to data, privacy and data, system security, and accessibility. We incorporate these requirements into a canvas model, the ethical requirements canvas. The canvas model serves as a practical business case tool in management practices. It not only facilitates the inclusion of ethical considerations but also highlights their business value, aiding management in understanding and discussing their significance in AI-enhanced environments.

Keywords: AI ethics · artificial intelligence · ethical requirements · IEEE Std 7000–2021 · ethical requirements canvas · software engineering

1 Introduction

The increasing integration of artificial intelligence (AI) into software engineering (SE) businesses is revolutionizing technology development, necessitating the incorporation of ethical requirements into management practices. This shift is emphasized by research [12, 30] and calls for aligning AI functionalities with ethical principles essential for guiding decision-making toward the development of trustworthy AI systems. Ethical requirements help to provide tangible actions derived from broader ethical principles like transparency, fairness, and privacy. For instance, the general principle of transparency becomes the need for “explainability” in AI, ensuring decision-making processes are clear and comprehensible for users [18]. As AI becomes more prevalent in sensitive sectors like health-care and education, SE organizations face increasing pressure from stakeholders, including developers, users, and regulators, to ensure AI systems like ChatGPT are not only innovative but also responsible and trustworthy [18, 30].

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 231–246, 2024.

https://doi.org/10.1007/978-3-031-53227-6_17

Creating AI systems that are ethical and in sync with societal norms is a crucial aspect of trustworthy AI [12, 29]. Despite this, SE management stakeholders who guide decision-making find it challenging to incorporate ethical requirements into their practices effectively [1, 5, 12]. A primary challenge lies in these stakeholders' determination of ethical requirements relevant to business and representing them accordingly in their management approaches [1, 5]. This difficulty is compounded by a noticeable disconnect among these stakeholders in recognizing the value of ethical requirements [1, 5]. Existing ethical guidelines further exacerbate this gap, primarily focused on the technical aspects of SE projects, often neglecting the equally critical managerial dimensions that guide decision-making [25, 36]. This omission leads to the undervaluation of ethical considerations and puts organizations at risk of legal, reputational, and regulatory repercussions [1, 4].

To address the challenge faced by SE management stakeholders in determining and valuing ethical requirements in AI systems, our study utilizes the IEEE Standard Model Process for Addressing Ethical Concerns during System Design (IEEE Std 7000–2021) [19]. This standard serves as a vital tool for concept exploration and the development of the concept of operations (ConOps) stage, offering a comprehensive roadmap for embedding ethical considerations in the creation and operation of autonomous and intelligent systems (A/IS). It encourages managerial stakeholders to actively engage in four critical areas: *Identifying* relevant ethical requirements for their System of Interest (SOI), *Eliciting* these requirements based on applicability, *Prioritizing* their importance, and *Incorporating* them into management strategies, considering key stakeholder success factors. While the standard acknowledges that ethical consideration is not solely the responsibility of management, it underscores the pivotal role of management in establishing ethical benchmarks and supervising their outcomes. Consequently, our research is driven by two fundamental questions:

RQ1: *What ethical requirements do SE management stakeholders consider crucial for AI-empowered SOI?*; and **RQ2:** *How can ethical requirements be effectively evaluated and integrated as success factors in SE management strategies for AI-empowered SOI?*

The primary aim of this study is to underscore the crucial role of ethical requirements for SE businesses, particularly in AI-enhanced environments. By addressing the outlined research questions, we seek to guide organizations to circumvent ethical pitfalls and cultivate a culture of trustworthiness in AI development. Our objective is to contribute significantly to the ongoing conversation about integrating ethics into AI and SE practices, ultimately aiming to bolster stakeholder trust and position organizations as frontrunners in ethical AI deployment.

The remainder of this study is organized as follows: Sect. 2 provides an overview of the background and existing literature, while Sect. 3 describes our research methodology, including data collection, analysis, and key findings. Discussions based on our insights are presented in Sect. 4, and Sect. 5 offers the study's conclusions.

2 Background

AI ethics aims to ensure AI technologies are developed and utilized in alignment with ethical and societal values, preventing unforeseen consequences or damage. It examines the ethical principles and moral concerns tied to the creation, implementation, and usage of AI systems [26]. While AI ethics encompasses worries about machine behaviors and the potential emergence of singularity intelligent AI [26], this study doesn't explore that dimension. Issues like bias, surveillance, job displacement, transparency, safety, existential threats, and weaponized AI underscore the imperative of instilling ethical considerations into AI engineering. Consequently, private, public, and governmental stakeholders have set AI principles as ethical guidelines. Notable among these are the EU's trustworthy AI guidelines (AI HLEG), IEEE's Ethically Aligned Design (EAD), the Asilomar AI Principles, and the Montreal Declaration for Responsible AI [18,19]. Guiding principles distilled from various guidelines, as outlined by Ryan and Stahl [32] and Jobin et al. [21], include Transparency, Justice, Non-maleficence, Responsibility, Privacy, Beneficence, Autonomy, Trust, Sustainability, Dignity, and Solidarity.

2.1 Ethical Requirements

Ethical requirements are multifaceted, requiring careful consideration and interdisciplinary collaboration spanning technology, law, philosophy, and social sciences [24]. Ethical requirements of AI are primarily from foundational ethical principles or rules, such as transparency and fairness, and are pivotal for fostering trustworthy AI [15]. They help interpret the guiding principles and standards that ensure AI systems' ethical design, creation, deployment, and operation. From the principle of privacy, for instance, an ethical requirement is privacy and data protection, entailing that AI systems should handle personal and sensitive data carefully according to legal regulations and best practices [15,21]. As such, they help build trust and align AI endeavors with human values and societal aspirations [15]. However, in SE, ethical requirements are predominantly articulated as functional and non-functional requirements during the development phase [15], yet they are seldom addressed at the management level, typically only insofar as to meet legal mandates like the General Data Protection Regulation (GDPR) [1,24].

2.2 Trustworthy AI

With the increasing integration of AI across various aspects of human life, the concept of Trustworthy AI has evolved to encompass a broader range of societal and environmental considerations. These include the implications for employment, societal equity, and the environment. Despite the presence of specific frameworks and guidelines from organizations, governments, and international bodies, the critical requirements that truly define what makes AI trustworthy remain a central concern [12,29]. The AI HLEG and IEEE EAD have been

instrumental in identifying critical ethical requirements, significantly shaping the discourse on trustworthy AI [18,19]. These frameworks outline key ethical principles that serve as a guide for both academia and industry professionals. The AI HLEG highlights seven key requirements for trustworthy AI: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability. Concurrently, the IEEE EAD emphasizes five: human rights, well-being, accountability, transparency, and awareness of AI's potential for misuse [19]. There's notable convergence in these requirements, which we explain as follows: *Human agency and oversight*: Emphasizes the importance of human rights and underscores the indispensability of human direction and supervision. *Technical robustness and safety*: Stresses the importance of crafting AI systems that resist threats, prioritize safety, have inherent protective mechanisms, and exhibit consistent, dependable, and replicable outcomes. *Privacy and Data Governance*: Navigates the privacy terrain, advocating the cause of data integrity, quality, and accessibility. *Transparency*: Entails a commitment to traceability, explainability, and effective communication of AI processes. *Diversity, non-discrimination, and fairness*: Encourages equitable AI practices, advocating for unbiased algorithms, universal design principles, and inclusive stakeholder engagement. *Societal and environmental well-being*: Focuses on AI's societal imprint, ranging from its ecological footprint to its broader societal repercussions and democratic implications. *Accountability*: Encompasses regularized auditing, transparent reporting, harm minimization, and effective remedial mechanisms. These enumerated requirements find application in tools like ECCOLA and Ethical User Stories (EUS), pivotal in executing the IEEE Std 7000–2021 approach of this study.

ECCOLA is an Agile-oriented method designed to enhance awareness and execution of AI ethics for developers in SE [36]. It synthesizes ethical requirements from AI HLEG and EAD, consolidating them into seven core themes or requirements and sub-requirements. The ECCOLA approach is a 21-card deck organized around seven primary requirements: transparency, data agency and oversight, safety and security, fairness, well-being, and accountability, and a stakeholder analysis card. Each requirement is represented further by one to six dedicated sub-requirement cards. ECCOLA is segmented into three components: the rationale behind its importance, actionable recommendations, and a tangible real-world example [36]. For direct access to ECCOLA, click here.

Ethical User Story concept integrates the user story methodology with an ethical toolset, facilitating the extraction of ethical requirements during technological design or development processes [16]. In SE and Agile methodologies, user stories help bridge business objectives and development activities by succinctly capturing customer demands [10]. These stories act as conduits to foster understanding between developers and users. They distill intricate concepts into more targeted information pieces, bolstering communication and collaboration

to ensure goal alignment. A standard user story is structured as: “As a [user role], I want [goal or need] so that [reason or benefit].” Here, the “user role” delineates a specific user’s identity or function. The “goal or need” specifies the desired outcome from the software, while the “reason or benefit” pinpoints the underlying motivation or value that drives this desire helping to concisely and clearly describe a user’s requirement for the SOI [10].

2.3 Standard Model Process for Addressing Ethical Concerns During System Design

The IEEE Std 7000–2021 provides a practical approach for SE businesses to identify and address ethical issues during the system design of their system of interest (SOI). We focus on the concept exploration and development of the concept of operations (ConOps) stage in our study, which emphasizes proactive communication with stakeholders, to help identify and prioritize ethical values to be integrated at the system design stage [20]. The procedure entails discerning these values from the operational concept, which lays out the system’s functionality, and from the value propositions and dispositions, which highlight the system’s benefits and potential outcomes. Central to the IEEE Std 7000–2021 are the Ethical Value Requirements (EVRs) concept. EVRs epitomize the essential worth of ethical requirements, ensuring that systems resonate with societal standards and uphold human rights, dignity, and well-being [12, 18, 20]. The standard advocates for meaningful engagement of primary stakeholders, especially those in management roles, throughout the design phase in *Identifying* pertinent ethical requirements by scrutinizing relevant ethical regulations, policies, and guidelines, including gathering stakeholder feedback. - *Eliciting* these ethical requirements based on their relevance to the SOI. - *Prioritizing* the inherent value of these requirements. - *incorporating* these values into the system’s core objectives and ensuring consistent communication and compliance monitoring with all concerned parties. Defining and embedding ethical requirements can bolster SOIs’ credibility, trustworthiness, and perceived value to help weave them seamlessly into their system’s design and development [20].

2.4 Implementing Ethical Requirements in SE Management

Aligning software development with an organization’s objectives is primarily achieved through SE management, which integrates critical success factors into operational and decision-making frameworks [14, 28]. Despite its importance, there’s a scarcity of tools that embed ethical requirements within SE management [3, 5]. Notably, the adaptation of canvas models for ethical representation is gaining traction among researchers and practitioners seeking to elevate ethical considerations in their practices [22, 27, 37]. Canvas tools are graphical representations that clarify intricate business concepts, facilitating stakeholder alignment. They break down various business facets, like customer segments or value propositions, into an easily digestible format often serving as a business snapshot enhancing understanding and communication [8, 28]. Some notable approaches

for the canvas model include The Ethics Canvas [22] which leverages the foundational blocks of the business model canvas to stimulate discussions on the ethical implications of technology. However, its scope on ethics is extensive and doesn't precisely target AI ethics or its requirements. The Open Data Institute's Data Ethics Canvas [27] offers a lens through which data practices can be ethically evaluated. Vidgen et al. [37] introduce a business ethics canvas, drawing inspiration from the applied ethics principles of the Markkula Center, which focuses on addressing data-centric ethical issues in business analytics. The canvas, however, predominantly focuses on the data ethics dimension. A more comprehensive canvas approach is the Trustworthy AI Implementation (TAII) canvas [2], which extends from the TAII framework [3]. It outlines the interplay of ethics within a company's broader ecosystem, touching upon corporate values, business strategies, and overarching principles but does not precisely pinpoint ethical requirements, potentially making it challenging for SE management stakeholders to translate it into actionable management practices [3].

3 Research Methodology

We adopt an exploratory approach to address our research questions. This approach is in line with Hevner et al.'s Design Science method, particularly the "build" component, given the innovative nature of our study and the limited resources in existing literature [17]. Exploratory methods provide valuable flexibility, especially when delving into less-explored research areas [35]. Hevner et al. emphasize the importance of adapting their seven guidelines, and our primary focus lies in developing conceptual artifacts, as outlined in their "Design as an artifact" guideline. While this phase typically yields conceptual insights rather than fully developed systems, the design science approach is crucial for shaping novel artifacts, even in the face of challenges [17].

3.1 Data Collection

We collaborated with 12 Finnish SE executives on an AI-enhanced project focused on autonomous marine transport for emission reduction and the enhancement of passenger and cargo experiences at the concept exploration stage. These executives represent various businesses specializing in different aspects of intelligent and autonomous SE, as detailed in Table 1. Our objective was to identify the essential ethical requirements these stakeholders deemed necessary for the AI-enabled System of Interest (SOI). To initiate our study, we secured the informed consent of our industry partners, emphasizing their entitlement to withdraw or request data deletion at any phase. Leveraging their SE background, which granted them a foundational understanding of the concepts, we embarked on a collaborative project segmented into three specific use cases. A series of workshops grounded on the brainstorming technique delineated by [33] facilitated the familiarization process with critical frameworks, including IEEE Std 7000-2021, ECCOLA, and the EUS concept.

During these sessions, the participants, who were predominantly executives, actively engaged in selecting pertinent ethical requirements from the 21 ECCOLA cards, highlighting those that resonated significantly with their business operations. The focus coalesced around ethical themes encapsulated by cards # 2 Explainability, # 3 Communication, # 5 Traceability, # 7 Privacy and Data, # 8 Data Quality, # 9 Access to Data, #12 System Security, # 13 System Safety, # 14 Accessibility, # 16 Environmental Impact, and # 18 Auditability. This careful selection served as a guide to pinpoint the ethical themes critical to their enterprise, facilitating a nuanced exploration. Extensive notes were documented to address subsequent inquiries and emerging concerns.

Table 1. SE Management Stakeholders

Solution provider	Area of expertise
Solution provider 1	Animation
Solution provider 2	Software development
Solution provider 3	Intelligent logistics
Solution provider 4	Remote and autonomous solutions
Solution provider 5	Transportation logistics
Solution provider 6	Computer controlled machinery
Solution provider 7	Intelligent translations
Solution provider 8	Intelligent transport infrastructure and logistics
Solution provider 9	Intelligent logistics
Solution provider 10	Information solutions
Solution provider 11	Automation solutions
Solution provider 12	Intelligent logistics

In eight workshops, each spanning one to three hours, we collaboratively formulated EUS using the ECCOLA method, tailoring the selections from ECCOLA to suit the requirements of each specific use case. Our detailed notes amounted to a total of 367, resulting in the creation of 253 EUS instances [34]. Examples of these instances include:

“As a [company CEO], with automated truck deliveries, I want [to have information, before sending my trucks on how data is handled], so that [I can feel secure that my data will not leak to unwanted parties].”

“As a [company data protection manager], I want to [authenticate the collected data] so that I can [ensure validity].”

“As a [system administrator], I want to [streamline the management of GDPR requirements] so that I can [ensure that the service remains unaffected by user information or data erasure requests].”

“As a [project stakeholder], I want the system [to feature clear and explainable logic] to [prevent project overruns or operational errors caused by unclear system descriptions].”

3.2 Data Analysis

We conducted our analysis utilizing content analysis, a systematic approach for dissecting qualitative data to discern recurring themes, patterns, and categories, ultimately yielding valuable insights [39]. In analyzing the EUS, we adopted an interpretive content analysis approach, prioritizing narrative interpretations of meaning over purely statistical inferences. This method enabled us to differentiate between manifest content, which represents overt messages in communication, and latent content, which encompasses subtle or underlying implications [39]. To streamline the analysis, we established a coding system. For instance, ‘TR’ was used as a code to symbolize ‘transparency’, while ‘DA’ represented ‘data’. These are just some examples of the various codes we employed throughout our analysis. These codes were then used to highlight specific ethical requirements within the dataset. For example, ‘TR’ pinpointed instances where transparency was a focal point in user stories. As we observed emerging patterns, we sought to identify correlations between the codes and overarching themes. These themes were then cross-referenced with central themes from the ECCOLA cards.

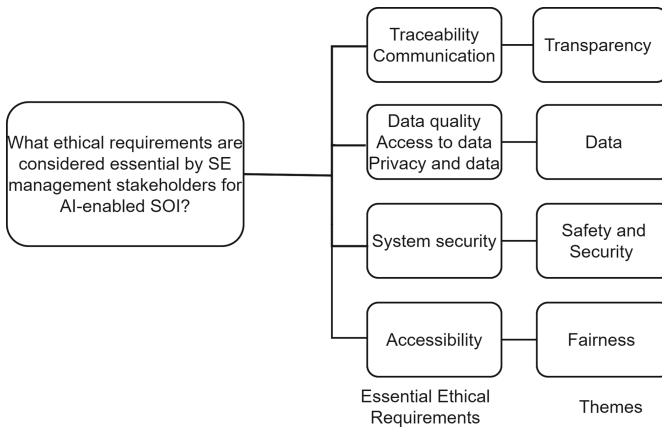
Utilizing the MoSCoW Prioritization technique [11], a popular tool in project management, software development, and business analysis, the executives classified the EUS based on their significance of “Must have, Should have, Could have, and Won’t have”. “Must have” captures indispensable requirements without which the project is incomplete. “Should have” comprised valuable yet non-critical elements; their omission wouldn’t jeopardize the project. “Could have” entails requirements that, while beneficial, aren’t urgent and can be tackled if resources permit. “Won’t have” covers those that are either irrelevant to the current project or simply unfeasible, possibly deferring them for later consideration or omitting them altogether [11]. The comprehensive prioritization can be found in Table 2. Of the 12 industry partners, nine participated in these classification exercises, while three were unavailable (denoted as N/A). The activity spanned several sessions, resulting in 177 out of the 253 EUS receiving priority rankings.

3.3 Findings

The prioritization from the EUS yielded seven distinct sub-requirements, categorized under four primary requirements. These sub-requirements are #5 Traceability, #3 Communication, #8 Data quality, #9 Access to data, #7 Privacy and data, #12 System security, and #14 Accessibility. They fall under the broader categories of Transparency, Data, Safety and Security, and Fairness. These emerged as crucial for SE management stakeholders, as illustrated in Fig. 1.

Table 2. Prioritization breakdown

Solution provider	Themes	Sub-Requirement	Prioritization
Solution provider 7	Transparency	#3 Communication	18
Solution provider 2	Transparency	#5 Traceability	42
Solution provider 10	Data	#7 Privacy and Data	20
Solution provider 9	Data	#7 Privacy and Data	29
Solution provider 12	Data	#8 Data Quality	7
Solution provider 4	Data	#8 Data Quality	7
Solution provider 6	Data	#9 Access to Data	3
Solution provider 2	Data	#9 Access to Data	5
Solution provider 10	Data	#9 Access to Data	4
Solution provider 11	Data	#9 Access to Data	3
Solution provider 12	Data	#9 Access to Data	3
Solution provider 8	Data	#9 Access to Data	2
Solution provider 9	Safety & Security	#12 System Security	24
Solution provider 7	Fairness	#14 Accessibility	10
Solution provider 1	N/A	N/A	-
Solution provider 3	N/A	N/A	-
Solution provider 5	N/A	N/A	-
Sum total of EUS	-	-	177

**Fig. 1.** Essential Ethical Requirements

4 Discussion

We examine our findings within existing research.

4.1 Essential Ethical Requirements

We analyze the seven identified ethical requirements and explore their significance and implications for stakeholders in SE management.

Traceability is pivotal in enhancing transparency and ensuring accountability within AI systems. It provides stakeholders with vital information to scrutinize and interpret the system's decisions [36]. By prioritizing traceability, those in SE management roles can effectively identify and manage the inherent risks associated with AI technology. This focus requires a detailed documentation process encompassing data sources, applied algorithms, computational models, and justifying particular outputs. Such comprehensive records identify potential weak points that could be prone to errors or biases, thereby enabling risk mitigation strategies to be deployed proactively [21]. As Ryan et al. underscore [32], maintaining stringent traceability practices reinforces accountability and fortifies customer and stakeholder trust, consequently elevating the organization's reputation.

Communication is central to disseminating essential details about an AI system's architecture, development phases, and functionalities to all pertinent stakeholders. Effective communication involves transparently articulating the system's objectives, capabilities, limitations, and possible repercussions. By doing so, stakeholders engaged in the project can gain a well-rounded understanding of the initiative's scope and aims, allowing them to identify and proactively address technical and ethical challenges. Open and transparent dialogue among SE management stakeholders can facilitate collaborative problem-solving and mitigate potential adverse outcomes. One challenge in communication within SE management is the complexity of technical jargon and the volume of information related to AI project documentation. However, prioritizing strategic communication can align expectations and clarify objectives [32].

Data Quality ensures that data serves its designated purpose and can be relied upon for making well-informed decisions within AI systems [6, 18, 23]. For SE management, data quality is a strategic component that influences the efficacy and efficiency of AI deployments. Subpar data quality elevates risks such as data breaches, security lapses, and other data-centric complications. These issues can inflate development expenses by necessitating the resolution of data inconsistencies, which in turn may lead to project delays and increased rework costs. Such disruptions can compromise the quality of AI solutions, diminishing customer satisfaction and eroding revenue and market share. Conversely, a commitment to high-quality data practices can assist SE management in curbing development costs, elevating product quality, enriching customer experience, and mitigating risks [18, 23].

Access to Data facilitates SE management by granting stakeholders insights into the data utilized in projects, development progression, and other pertinent details, aiding in identifying and mitigating risks associated with their chosen data for SOI. As businesses accumulate vast and diverse data sets, maintaining streamlined access becomes indispensable to prevent data landscapes from turning chaotic and complex [3]. Moreover, with tightening regulatory landscapes, such as the GDPR and the California Consumer Privacy Act (CCPA), adept data management, particularly regarding access, has gained paramount significance. Conversely, inefficient practices regarding data access can result in gaps in understanding data’s availability, quality, security measures, proprietorship, and overarching governance [18].

Privacy and Data are key elements in maintaining the integrity of AI systems, safeguarding against data breaches, and avoiding biased or discriminatory outcomes. AI systems often require access to data, including sensitive or personal information, that demands stringent protection measures. SE management stakeholders can play a vital role by incorporating strong privacy and data handling practices. These measures enable the ethical utilization of data, safeguarding against biased or prejudicial data sets and avoiding harm to individuals or groups. Wang et al. [38] point out that while data can provide invaluable benefits to organizations, it can also pose risks. High-profile cases like Meta (formerly Facebook) underscore the necessity for striking a balanced approach between exploiting data’s benefits and mitigating its associated risks, both from a social and regulatory standpoint.

System Security focuses on deploying security protocols like authentication and encryption to safeguard against unauthorized system or data access while ensuring that the system can quickly recover from any security breaches. The ultimate objective is to guarantee the system’s safe and reliable operation across diverse scenarios without harming users or society. Cheatham et al. [9] note that AI technology’s relative infancy means that SE management stakeholders often lack the refined understanding necessary to grasp societal, organizational, and individual risks fully. This lack of understanding can lead to underestimating potential dangers, overvaluing an organization’s ability to manage those risks, or mistakenly equating AI-specific risks with general software risks. To avoid or minimize unforeseen consequences, these stakeholders must enhance their expertise in AI-related risks and involve the entire organization in comprehending both the opportunities and responsibilities of AI technology.”

Fairness entails management practices of avoiding biased algorithms or data sets that may lead to discrimination or unfair treatment of certain groups [18]. It also means ensuring that AI systems design and development are supervised not to perpetuate or exacerbate societal inequalities. Berente et al. [5] explain that management stakeholders can ensure that the teams responsible for developing

and deploying AI systems are diverse regarding gender, race, and ethnicity to mitigate bias in decision-making. Diversity can help ensure that AI is designed and deployed fairly and ethically for all users, thereby increasing the adoption and acceptance of AI by a broader range of users.

4.2 Towards a Business Case for Ethical Requirements

To address RQ2 effectively, we introduce the Ethical Requirements Canvas, depicted in Fig. 2. This canvas serves to underline not just the importance but also the intrinsic value of ethical requirements, thereby constructing a business case for their integration. Business cases are essential for management to evaluate a project’s costs, benefits, risks, and alternatives, ensuring alignment with the organization’s strategic goals [40]. The Ethical Requirements Canvas serves as a practical instrument that not only integrates ethical considerations into management practices but also highlights their business value [28]. Consequently, the canvas provides a pragmatic method for aligning ethical requirements with the organization’s broader goals, articulating their significance and potential for adding value in business terms.

Ethics Requirements Canvas			Project Title	
Ethical Requirements	Key activities	Value proposition	Impact	Stakeholders
<i>Traceability</i>	<i>What activities are needed for implementation?</i>	<i>Promotion of Trustworthy AI practices</i>	<i>How does it impact your employees and customers?</i>	<i>Identify the categories of stakeholders that can be affected</i>
<i>Communication</i>				
<i>Data quality</i>	3	<i>Enhance stakeholder engagement and buy-in</i>	5	
<i>Access to data</i>	Key Resources	<i>Increased trust in AI technologies among users and society</i>	Societal Impact	
<i>Privacy and data</i>			<i>How does it impact society at large?</i>	
<i>System security</i>				
<i>Fairness</i>	<i>What resources are needed for implementation?</i>	<i>Increased sales</i>	6	2
1	4	9	6	2
Cost			Benefits	
<i>What are potential negative impact that can result from a lack of implementation in our products or services such as reputational and financial impacts?</i>			Value stream	
<i>Discuss the negative impacts such as legal, privacy, environmental impacts, employment impacts etc.</i>			<i>Will customers need this and are they willing to pay?</i>	
<i>Redress costs from AI systems failing to operate or to be used as intended?</i>			<i>Which aspects of can be monetized?</i>	
<i>Adapted from The Business Model Canvas (Osterwalder, A., & Pigneur, Y. (2010))</i>			<i>Are customers willing to pay for it over what is obtainable</i>	
7			8	

Fig. 2. Ethical Requirements Canvas

Section one presents the ethical requirements identified through our research. It’s important to note that these requirements are displayed for reference and awareness, not for rigid adherence. Section two focuses on identifying the organization’s stakeholders. Here, SE management can discuss various categories of

stakeholders, such as human and non-human agents, different age groups, societal standing, and levels of vulnerability, among others. Section three outlines the essential business operations necessary to realize the value proposition of integrating ethical requirements. Section four lists the resources required for effective implementation. Sections five and six allow SE management to assess the societal, internal, and external impacts of incorporating these ethical parameters into their SOI. Section seven explores the financial, reputational, or otherwise costs associated with choosing to integrate or overlooking ethical requirements. Section eight evaluates the benefits and potential monetization of ethical requirements. Section nine illuminates the distinct advantages of ethical considerations, assisting in identifying vital initiatives that enhance the benefits of ethical requirements, potentially serving as critical determinants of success [7]. These benefits encompass elevating the organization to a Trustworthy AI business status, akin to the positive reputational impact observed in companies with sustainability initiatives. This can enhance stakeholder engagement-from the business being perceived as ethical and trustworthy-and potentially expanding market share and boosting profitability due to increased user trust. [7,27,28].

While the Ethical Requirements Canvas provides a systematic framework for visualizing and assessing ethical considerations, it may have inherent limitations. Its structured nature could risk simplifying complex ethical dilemmas, potentially fostering a compliance-centric mindset at the expense of cultivating a deeper ethical culture [31]. This approach risks satisfying only the minimum legal standards rather than aspiring to ethical excellence, which may lead to the marginalization of crucial ethical aspects [13,28,31]. Additionally, while adaptability is one of the canvas's strengths, it also poses challenges. Our research identified seven core ethical requirements, but their relevance and prioritization can differ significantly among organizations due to unique contextual factors, industry norms, and stakeholder expectations. Therefore, it is critical to balance adherence to industry standards with the strategic objectives of the organization when applying the canvas.

4.3 Limitation

A limitation inherent to our research is its specific focus on the marine transportation sector within Finland, potentially circumscribing the external validity and generalizability of our findings to other geographical contexts or industries experiencing AI-driven digital transformations. Despite this, we argue that our research lays a foundational framework that can be adapted and scrutinized in various settings [33].

For future studies, we plan to validate the Ethical Requirements Canvas via workshops with SE management teams and industry-wide surveys. These evaluations will not only gauge the canvas's usability and relevance but will also fine-tune its alignment with both organizational demands and ethical standards.

5 Conclusion

In this study, we have made three principal contributions. First, we compiled a comprehensive set of ethical requirements reflecting the perspectives of SE management stakeholders. Second, we presented a stakeholder-centric approach that is responsive to the challenges faced by the industry. Third, we introduced the “Ethical Requirements Canvas,” a novel tool designed to elucidate and integrate the value of ethical considerations into SE management practices. The canvas not only acts as an ethical roadmap for practitioners but can also facilitate risk management and promote judicious decision-making [28]. From an academic standpoint, our framework lays the groundwork for further inquiry into the integration of ethical requirements in AI and SE management, encouraging cross-disciplinary research and assessments of tool efficacy. On a practical level, our work supports SE managers in embedding ethical principles more deeply within their processes, thereby advocating for the development of trustworthy AI systems.

References

1. Agbese, M., Mohanani, R., Khan, A., Abrahamsson, P.: Implementing AI ethics: making sense of the ethical requirements, EASE 2023, pp. 62–71. Association for Computing Machinery, New York, NY, USA (2023)
2. Baker-Brunnbauer, J.: Tair framework canvas (2021)
3. Baker-Brunnbauer, J.: TAI Framework, pp. 97–127. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-18275-4_7
4. Barn, B.S.: Do you own a Volkswagen? Values as non-functional requirements. In: Bogdan, C., et al. (eds.) HESSD/HCSE -2016. LNCS, vol. 9856, pp. 151–162. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44902-9_10
5. Berente, N., Gu, B., Recker, J., Santhanam, R.: Managing artificial intelligence. *MIS Q.* **45**, 1433–1450 (2021)
6. Bobrowski, M., Marr, M., Yankelevich, D.: A software engineering view of data quality, February 1970
7. Boehm, B.: Value-based software engineering: reinventing. *ACM SIGSOFT Softw. Eng. Notes* **28**(2), 3 (2003)
8. Carter, M., Carter, C.: The creative business model canvas. *Soc. Enterp. J.* **16**(2), 141–158 (2020)
9. Cheatham, B., Javanmardian, K., Samandari, H.: Confronting the risks of artificial intelligence. *McKinsey Q.* **2**(38), 1–9 (2019)
10. Cohn, M.: *User Stories Applied: for Agile Software Development*. Addison-Wesley Professional, Boston (2004)
11. Consortium, D.: DSDM project framework. <https://www.agilebusiness.org/dsdm-project-framework/moscow-prioritisation.html>, January 2014
12. Dignum, V.: *Responsible Artificial Intelligence: how to Develop and use AI in a Responsible Way*. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-30371-6>
13. Elliott, K., et al.: Towards an equitable digital society: artificial intelligence (AI) and corporate digital responsibility (DDR). *Society* **58**(3), 179–188 (2021)

14. Freeman, P.: Software engineering body of knowledge (SWEBOK). In: International Conference on Software Engineering, vol. 23, pp. 693–696 (2001)
15. Guizzardi, R., Amaral, G., Guizzardi, G., Mylopoulos, J.: Ethical requirements for AI systems. In: Goutte, C., Zhu, X. (eds.) Canadian AI 2020. LNCS (LNAI), vol. 12109, pp. 251–256. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-47358-7_24
16. Halme, E., et al.: How to write ethical user stories? Impacts of the ECCOLA method. In: Gregory, P., Lassenius, C., Wang, X., Kruchten, P. (eds.) XP 2021. LNBP, vol. 419, pp. 36–52. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78098-2_3
17. Hevner, A., March, S., Park, J., Ram, S.: Design science in information systems research. *MIS Q.* **28**(1), 75–105 (2004)
18. High-Level Expert Group on Artificial Intelligence (AI HLEG): Ethics guidelines for trustworthy AI. Technical Report, European Commission, April 2019
19. IEEE: Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems, version 2 (2017). https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf
20. IEEE: IEEE standard model process for addressing ethical concerns during system design. IEEE STD 7000–2021, pp. 1–82 (2021)
21. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019)
22. Lewis, D., Reijers, W., Pandit, H., Reijers, W.: Ethics canvas manual (2017)
23. Loshin, D.: The practitioner’s guide to data quality improvement. Elsevier (2010)
24. Mökander, J., Floridi, L.: Ethics-based auditing to develop trustworthy AI. *Mind. Mach.* **31**(2), 323–327 (2021)
25. Morley, J., Floridi, L., Kinsey, L., Elhalal, A.: From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* **26**(4), 2141–2168 (2020)
26. Müller, V.C.: Ethics of artificial intelligence and robotics (2020)
27. Open data institute: data ethics canvas (2019). <https://theodi.org/article/data-ethics-canvas/>
28. Osterwalder, A., Pigneur, Y.: Business Model Generation: a Handbook for Visionaries, Game Changers, and Challengers, vol. 1. Wiley, Hoboken (2010)
29. Papagiannidis, E., Enholm, I.M., Dremel, C., Mikalef, P., Krogstie, J.: Deploying AI governance practices: a revelatory case study. In: Dennehy, D., Griva, A., Pouloudi, N., Dwivedi, Y.K., Pappas, I., Mäntymäki, M. (eds.) I3E 2021. LNCS, vol. 12896, pp. 208–219. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-85447-8_19
30. Ray, P.P.: ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys. Syst.* **3**, 121–154 (2023)
31. Ricart, J., Casadesus-Masanell, R.: How to design a winning business model. *Harv. Bus. Rev.* **89**(1–2), 100–107 (2011)
32. Ryan, M., Stahl, B.C.: Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *J. Inf. Commun. Ethics Soc.* **19**(1), 61–86 (2020)
33. Shull, F., Singer, J., Sjøberg, D.I.: Guide to Advanced Empirical Software Engineering. Springer, London (2007). <https://doi.org/10.1007/978-1-84800-044-5>
34. Staron, M.: Action Research in Software Engineering. Springer, Cham (2020). <https://doi.org/10.1007/978-3-030-32610-4>

35. Swedberg, R.: Exploratory research. In: The production of knowledge: Enhancing Progress in Social Science, pp. 17–41 (2020)
36. Vakkuri, V., Kemell, K.K., Jantunen, M., Halme, E., Abrahamsson, P.: Eccola-a method for implementing ethically aligned AI systems. *J. Syst. Softw.* **182**, 111067 (2021)
37. Vidgen, R., Hindle, G., Randolph, I.: Exploring the ethical implications of business analytics with a business ethics canvas. *Eur. J. Oper. Res.* **281**(3), 491–501 (2020)
38. Wang, C., Zhang, N., Wang, C.: Managing privacy in the digital economy. *Fundam. Res.* **1**(5), 543–551 (2021)
39. Weber, R.P.: *Basic Content Analysis*, vol. 49. Sage, Thousand Oaks (1990)
40. Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B., Wesslén, A.: *Experimentation in Software Engineering*. Springer, Berlin (2012). <https://doi.org/10.1007/978-3-642-29044-2>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





What Is the Cost of AI Ethics? Initial Conceptual Framework and Empirical Insights

Kai-Kristian Kemell¹(✉)  and Ville Vakkuri² 

¹ Department of Computer Science, University of Helsinki, Pietari Kalmin Katu 5, 00560 Helsinki, Finland

kai-kristian.kemell@helsinki.fi

² School of Marketing and Communication, University of Vaasa, Wolffintie 32 FI-65200 Vaasa PL 700, 65101 Vaasa, Finland

ville.vakkuri@uwasa.fi

Abstract. AI ethics has become a common topic of discussion in both media and academic research. Companies are also increasingly interested in AI ethics, although there are still various challenges associated with bringing AI ethics into practice. Especially from a business point of view, AI ethics remains largely unexplored. The lack of established processes and practices for implementing AI ethics is an issue in this regard as well, as resource estimation is challenging if the process is fuzzy. In this paper, we begin tackling this issue by providing initial insights into the cost of AI ethics. Building on existing literature on software quality cost estimation, we draw parallels between the past state of quality in Software Engineering (SE) and the current state of AI ethics. Empirical examples are then utilized to showcase some elements of the cost of implementing AI ethics. While this paper provides an initial look into the cost of AI ethics and useful insights from comparisons to software quality, the practice of implementing AI ethics remains nascent, and, thus, a better empirical understanding of AI ethics is required going forward.

Keywords: Ethics · Machine learning · Cost estimation · Software engineering · Artificial intelligence

1 Introduction

Despite AI ethics being increasingly discussed both on the academia and now out on the field as well, it remains of secondary importance in practice [13, 15]. While companies are becoming aware of the potential importance of AI ethics, its practical implementation is still an on-going issue. In research, this continues to manifest as a lack of empirical studies on the topic. While some companies show interest towards AI ethics and even release statements about their commitment to developing ethical software systems, little is known how this is done in practice, given the lack of empirical studies on AI ethics [13].

As little is known about the practical implementation of AI ethics, it is also difficult for companies to evaluate the resources and costs required for doing

so. Indeed, especially from a business point of view, AI ethics remains an open question. While the potential benefits of implementing ethics are becoming more clear for software companies (through the potential cost of ignoring ethics, if nothing else), and few companies would go on record to say ethics is not a priority for them, the cost of AI ethics remains unclear.

Ethics encompasses the entirety of the development process, from design to operations. At different points of the process, ethics manifests in different ways in SE practice [16]. Early on, design decisions shape the system, and ethical issues can arise from major decisions such as the business logic or the very nature of the system [4]. During development, ethics includes issues from data to end-user involvement (e.g., as seen through the plethora of tools included in tool review of Morley et al. [10], and as highlighted by the ECCOLA method [16]). During operations, ethics may necessitate new metrics to monitor; there are some examples of issues in AI systems having recently been uncovered through bad publicity on social media (e.g., a chatbot giving unauthorized diet advice for users seeking help for eating disorders¹).

Ethics is more than just minimum compliance to laws and regulations. At worst, ignoring ethical issues can lead to a system being pulled from production. Because ethics encompasses the entire development process, fixing issues stemming from poor design decisions early on can be highly costly and difficult in production. The ease of fixing issues early on in the development process is an acknowledged phenomenon in software quality [11], as well as, arguably, software development overall.

In this paper, we provide an initial look at AI ethics from the point of view of business by (1) discussing its relevance for business, and (2) discussing it from the point of the resources needed for implementing ethics. It is established in extant literature that there are still prominent gaps to be addressed in the practical implementation of AI ethics, and the business and resource point of view is one of them. We build this discussion on both existing literature and data from three empirical cases. By utilizing existing literature on software quality, we propose a high-level cost framework for ethics in SE. Then, through the example cases, we provide some initial insights into what types of activities, and thus, costs, are associated with implementing ethics in practice in SE.

While this paper is specifically motivated by *AI ethics*, this discussion is relevant for ethics in SE overall. For example, issues such as green IT are a part of AI ethics but also relevant for software organizations overall. We have chosen AI ethics as the context for this paper due to its timeliness and due to nature of the data we have collected.

The rest of this paper is structured as follows. Section 2 presents the theoretical background of the paper by discussing existing literature. In Sect. 3, we discuss the cost of (AI) ethics by building on existing literature on software quality and utilizing an existing cost framework for software quality. In Sect. 4, we provide some initial insights into the cost of (AI) ethics by utilizing past data we have originally collected for other research purposes (specifically, to develop

¹ <https://www.theguardian.com/technology/2023/may/31/eating-disorder-hotline-union-ai-chatbot-harm>.

the ECCOLA method [16]). In Sect. 5 we discuss the theoretical and practical implications of this paper. Section 6 concludes the paper.

2 What and Why Ethics

In Sect. 2.1, we provide a general overview of ethics in relation to SE, and more specifically AI. In Sect. 2.2, we expand on this discussion by adding a business focus.

2.1 Ethics, Ethics in SE, and AI Ethics

Ethics can be described as a philosophical field of study. In particular, ethics is the study of morality. In this paper, we discuss *applied ethics*, specifically in the context of both business ethics and ethics in SE, and more specifically, AI ethics. Applied ethics examines real-life situations, which are often unclear or debatable, in order to understand what would be the right or wrong action to take with the given set of values. E.g., why should software companies care about the environment (green IT)? Additionally, applied ethics can be thought of as 'ethics as practice' [1, 18], examples of which are guidelines and codes of conduct in SE or AI ethics.

The current discussion on AI ethics stems from the tradition of computer ethics where ethical discussion includes the ethics of system development and use, among other topics (see, e.g., [7]). Over the decades, this discussion has included topics such as piracy, green IT, cybersecurity, automatization and, more recently, AI ethics. The current discussion on AI ethics also draws from the various past discussions on ethics in SE, including topics such as business and the societal impacts of IT.

AI ethics is often approached through principles. Jobin et al. [8], based on their extensive review of AI ethics guidelines, outline the most commonly discussed principles: transparency, justice and fairness, non-maleficence, responsibility, and privacy. For example, fairness deals with issues related to bias and discrimination, which manifest in practice as, e.g., issues in ML system outputs and training data. However, bringing these principles into practice remains an on-going challenge in the area, as the guidelines seem to not have had a notable impact on industrial practice [13, 17] based on empirical studies, supporting the argument of Mittelstadt [9] about the ineffectiveness of principles alone. In fact, the practical implementation of AI ethics in general remains a topical challenge in ML development, and empirical studies remain scarce [10, 12]. While in addition to numerous conceptual papers, a number of papers discussing the technical implementation of, e.g., fairness (Fairness 360 etc.) exist, reported industry use cases and empirical studies are lacking.

2.2 Why (AI) Ethics?

While some organizations may still be pondering the business relevance of ethics, especially in the field of AI, ethics has gained mainstream attention. Ethical failures and potential ethical issues have been extensively discussed in mainstream

media, and companies developing ML solutions have attempted to react to this discussion by, for example, publishing their own guidelines for AI ethics (see Jobin et al. [8]) in order to signal commitment to the values within. Though good or bad publicity is a large motivator for companies to consider AI ethics, there are arguably various potential benefits for doing so. These (may) include: (1) brand equity² (2) consumer adoption, (3) social acceptance, (4) employee satisfaction³ (5) investor relations (ESG reporting), (6) market entry requirements (EU GDPR; upcoming EU AI Act), (7) proactive approach to laws and regulations (e.g., upcoming EU AI Act), (8) avoiding costly changes in production [16], and (9) a systematic approach to ethics over an ad hoc one [16].

Brand equity refers to good or bad publicity. There have been various ethical failures that have made the news, resulting in bad publicity and typically necessitating actions taken to correct the situation. Similarly, *consumer adoption* can be negatively impacted by ethical issues. Users are becoming increasingly conscious about issues such as data privacy and fairness, and tackling such topics in an ethical manner can become a selling point in ML. In a more general sense, *social acceptance* becomes important when developing particularly disruptive technologies that impact society or an organization on a larger scale, outside the scope of just their users. For example, autonomous vehicles impact traffic as a whole, rather than just their passengers (“drivers”). Aside from external stakeholders, consideration of AI ethics can also improve *employee satisfaction* in a similar manner to improving consumer opinion. If your values strongly conflict with those of your employees, it may lead to conflicts or resignations. Moreover, *investor relations* (ESG: Ecological, Social, and Governance), can also be improved via attention to AI ethics.

Market entry requirements, in this case, refers to the relevant laws and regulations. In particular, the European Union with its GDPR and the upcoming AI Act that are directly related to AI ethics, may necessitate more ethical consideration than the local region of the company. To this end, AI ethics can foster a *proactive approach to laws and regulations* can help companies adapt to the changing regulatory landscape for ML systems, with new regulations and laws constantly discussed (e.g., recently for Large Language Models (LLMs) and Generative AI) across the globe.

Ethics, like quality [11], is arguably easier to implement early on in software development, and thus, doing so can help in *avoiding costly changes in production*. Ethics encompasses the entire development process from design to production [16]. Finally, by actively pursuing AI ethics, companies are able to utilize a *systematic approach to ethics over an ad hoc one*. Even when ethics is not implemented actively, values still make their way into the product nonetheless.

² E.g., the capital of Finland, Helsinki, advertising their commitment to ethical AI: <https://www.hel.fi/fi/uutiset/helsinki-laati-periaatteet-datan-ja-tekoalyn-eettiselle-kaytolle>,.

³ E.g., <https://www.wired.com/story/google-brain-ai-researcher-fired-tension/>,.

3 Research Framework: Cost of Quality, and the Relationship of Quality and Ethics

In this section, we present and justify our approach to discussing the cost of AI ethics. We make a comparison to quality, which, as we argue in Sect. 3.1, shares some (historical) similarities with the current state of (AI) ethics. In Sect. 3.2, based on existing literature, we present an overview of the types of costs associated with quality and, building on it, propose a similar cost structure for AI ethics.

3.1 Is Ethics Just Another Quality Feature?

We argue that we are currently seeing various parallels between the current state of AI ethics and the historical evolution of software quality assurance. Software quality was, in the past, often overlooked in favor of more immediate business concerns such as time-to-market or simple profitability. Over time, it evolved to be an integral and integrated part of the SE process. To some extent, we currently are seeing similar developments in AI ethics. Despite the discussion on the growing importance of AI ethics, it is still typically largely overlooked in practice [13, 15]. Though companies are increasingly becoming aware of ethics-related issues such as fairness, the industry still seems to lack systematic frameworks and processes for implementing AI ethics, or at least it fails to utilize them.

In this paper, we approach ethics from the point of view of quality, to provide a point of comparison with an existing, well-established phenomenon in SE. While ethics is *not* simply quality and the two are not fully analogous, we nonetheless make this comparison due to the various similarities they do share:

- *Overlooked importance.* Historically, software quality was seen as a secondary objective, much like ethics currently. Its importance was acknowledged after initial failures, but making it a part of SE practice took its time. This has also been the case in AI ethics, with its importance largely now acknowledged but its practical implementation still a challenge [13].
- *Long-term consequences for software.* Both ethics and quality can result in severe negative impacts for the system(s) being developed if overlooked. Much like how bugs can render a system unusable, unforeseen ethical issues can result in an ML system being pulled from production (e.g., as was the case with the chatbot mentioned in Sect. 1).
- *Interdisciplinary nature.* Much like how Quality Assurance (QA) requires the involvement of various stakeholders other than just software developers, implementing AI ethics is also a multidisciplinary effort. While developers (and ML experts) are the ones bringing ethics into practice, the process still involves other stakeholders as well (e.g., ethics committee, experts, users...).
- *Maturing over time.* Software quality has evolved over time from simple debugging to formal QA processes and a continuous SE process (CI/CD). AI ethics seems to also be moving from a minimal regulatory and legal compliance to the development of ethical frameworks (e.g., ECCOLA [16]) and processes, although this is still on-going [14].

- *Relevance of organizational culture.* The implementation of ethics, like quality, is unlikely to succeed if it is an afterthought or a tacked-on process. AI ethics needs to become a part of organizational culture, and to this end, it needs to become a natural part of SE (e.g., as professional norms [6]).
- *Harder and costlier to implement in production.* Quality is cheaper to implement earlier on in the SE process [11]. This is also arguably the case for ethics as well. As we discuss next, ethics also encompasses system design and business logic. A system where the core (ethical) issue stems from the very goal of the system is difficult to fix in production, to say the least.

This comparison between ethics and quality is not a novel thought of ours. Existing literature has made similar observations. For example, in the literature review of Giray [5], AI ethics topics such as fairness are explicitly referred to as new types of *quality requirements* for ML systems. Indeed, it can be argued that, if quality is about assuring that the system works as intended, ethics shares the same goal on a conceptual level: assuring that the system works as intended (from the chosen ethical point of view).

However, AI ethics is not just software quality, especially not as it is conventionally understood. While some AI ethics principles such as *predictability*, which focuses on ensuring the system produces intended outputs or results reliably, are closely related to conventional software quality goals, AI ethics also encompasses system *design* and *business* in addition to software development [16]. A technically sound system that is of high quality can still be unethical. E.g., widespread AI-based surveillance using facial recognition is typically considered unethical as a concept (e.g., in the draft of upcoming AI act such systems are labelled as being of ‘unacceptable risk’) – and yet the use of such systems in contexts such as airport security would be considered acceptable by many, highlighting the complex nature of AI ethics.

As opposed to seeing (some parts of) ethics as quality issues, an argument could be made that it is in fact quality that is a part of ethics in SE. The ACM Code of Ethics and Professional Conduct discusses quality as a part of the job responsibilities of a software professional. It remarks that one should “strive to achieve high quality in both the processes and products of professional work” [6]. Regardless, this further provides justification for the parallels we draw between the two in the context of this paper.

3.2 The Cost of Ethics

Based on Sect. 3.1, we argue that quality offers a familiar point of reference (in SE) for initially approaching ethics from a cost point of view. According to Slaughter et al. [11], costs of quality consist, on a high level, of conformance and nonconformance. *Conformance* refers to the costs associated with developing quality products (i.e., ‘doing’ quality). *Nonconformance* refers to the costs resulting from failures resulting from poor quality (i.e., not ‘doing’ quality).

In more detail, Slaughter et al. [11] split the costs of conformance to prevention and appraisal costs. *Prevention costs* are associated with “preventing

defects before they happen”, which “include the costs of training staff in design, methodologies, quality improvement meetings, and software design reviews” [11]. *Appraisal costs*, on the other hand, include “measuring, evaluating, or auditing products to assure conformance to quality standards and performance. For software, examples of appraisal costs include code inspections, testing, and software measurement activities” [11].

Costs of nonconformance are further split into internal failure costs and external failure costs by Slaughter et al. [11]. *Internal failure* costs “occur before the product is shipped to the customer. For software these include the costs of rework in programming, reinspection, and retesting.” [11] *External failure* costs “arise from product failure at the customer site. For software, examples include field service and support, maintenance, liability damages, and litigation expenses.” [11]

In practice, from the point of view of the SE process, they assign these costs to three phases:

1. Software Quality Investment (SQI). The initial investment of doing quality. This includes “the initial expenses for training, tools, effort, and materials required to implement the quality initiative.” [11]
2. Software Quality Maintenance (SQM). Maintaining the processes set up during SQI. Ongoing expenditures “for meetings, tool upgrades, and training that are required to maintain the quality process.” [11]
3. Software Quality Revenues (SQR). Any resulting revenue. Revenues derived from “projected increases in sales or estimated cost savings due to the software quality improves.” [11]

Based on this, we propose a similar typology for the cost of AI ethics. We propose the following phases for AI ethics from a business point of view:

1. Ethics Investment. The initial investment for incorporating ethics into SE. This includes a wide variety of costs, such as: recruiting new experts, adopting new methods or other SE tools, modifying existing SE processes or creating new ones, more systematic project documentation, training, materials, etc.
2. Ethics Maintenance. Costs of maintaining the processes established in the first step. These include salaries of any new hired experts, meetings and other recurring tasks, etc.
3. Ethics Revenues. Any resulting revenue originating from the previous steps, such as increases in sales, brand equity, cost savings from failure prevention, etc.

Arguably, this is still a very nascent area of research. Because the practice of AI ethics overall is still poorly understood compared to software quality, the latter of which has decades of history of practice behind it by now, the associated processes are still being shaped out on the field. Thus, providing a comprehensive and detailed framework for the cost of AI ethics at this stage is not feasible. However, past simply proposing this typology on a conceptual level, we also

provide an initial look at the cost of AI ethics in practice. In Sect. 4, we focus especially on the first phase, the initial ethics investment, through empirical insights from three past cases we have worked on.

4 Empirical Examples

In this section, we use empirical data to provide an initial look at what types of processes are required to implement ethics and what kinds of activities result in costs when doing so. In Sect. 4.1, we describe the cases that the examples are from. In Sect. 4.2, based on these cases, we discuss the practicalities of implementing ethics from the point of view of resources and costs.

4.1 Cases and Data Description

To illustrate what the cost of implementing AI ethics means in practice, we build on three cases. Each case organization worked on a project where ethics was considered one of the key requirements. One of the projects was a blockchain project and the other two were ML development projects. The cases are illustrated in (table below 1)

Through these cases, we provide an initial look at the cost of implementing (AI) ethics, focusing on the initial *ethics investment*, as well as some early insights into *ethics maintenance* (Sect. 3.2). We utilize multiple types of data for each project, including interviews, project documentation, notes from workshops with developers, observation, etc. We feel that the use of a varied set of data lets us better explore a novel phenomenon such as this by giving us a clearer picture of what kinds of resources were needed to actively tackle ethics in a software development project. The types of data for each case are detailed in Table 1.

This data is used to illustrate what types of activities are associated with implementing AI ethics into practice, which are then discussed from the point of view of the types of costs discussed in Sect. 3.2. Thus, in terms of analysis, our focus is simply on *what* was done in the project to implement ethics, and what resources were needed to do so. As empirical studies in AI ethics are still lacking (see e.g. [10, 16]), our understanding of what types of processes are needed to do so is consequently lacking as well. Through these cases, we are able to provide an initial look at the cost of AI ethics by looking at what types of activities may be involved when implementing AI ethics in practice. These cases let us evaluate the feasibility of the framework before further data collection.

Moreover, in this paper and these three cases, we approach ethics through specific ethical frameworks, which vary by case. As the study of Jobin et al. [8] highlights, there is a lack of a clear understanding of what exactly AI ethics is, or should be, with different principles being used in different contexts to approach AI ethics. By utilizing existing ethical frameworks, we (and the case organizations, more importantly) are able to clearly define *what* ethics means in the context of each case. This important as it also helps define what an ethical system should look like, and thus helps define what actions should be taken to reach that goal, directly affecting *how* ethics is implemented in each case.

Table 1. Overview of cases and data.

#	Context	Data sources	Data types
1	Blockchain	1 developer	Interviews, project documentation, developer notes
2	ML (predicting tool)	1 development team	Project documentation, notes from workshops with developers
3	ML (voice recognition)	Client company & 1 development team	Project documentation, notes from workshops with developers

4.2 Case 1

Case 1 summary:

- **Project context:** Data from a single developer working in a research-industry collaboration blockchain project.
- **Who implemented ethics:** As the project progressed, involving ethics into the project became the responsibility of a single developer. The developer discussed matters with an external ethics expert as needed.
- **Ethical framework used:** EU Guidelines for Trustworthy AI [3] & Prototype of ECCOLA [16], which was being developed at the time.

Project activities related to ethics (time spent) [stakeholders involved] in case 1:

- Decision to implement ethics made in a design meeting (2h). [Project management and developers]
- Initial training with ethics expert (1h). [Ethics expert and developer]
- Ethics as a part of biweekly iteration planning (1–2h x n) [Developer and scrum master]
- Use of ethical tool during development (?h) [Developers]
- Additional ethics documentation (1–5 sheets per iteration) [Developer]
- Expert hotline (?h) [Developer and ethics expert]
- Internal presentations documenting the implementation of ethics in the project [Developer and project management]

Case 1 Observations. In case 1, we observed most resources spent on ethics being spent early on in the project (i.e., on **ethics investment**). As the project progressed, although ethics resulted in recurring resource investments (expert hotline; role in biweekly planning), the investment was largely frontloaded. Simply defining *what* the investment (i.e., ethics) is takes resources, as ethics in SE is a novel phenomenon that requires clarification in each project context.

In this regard, one challenge was the project context: the project as a blockchain project, and no ethical frameworks for that particular project context were identified at the time. As a result, frameworks for AI ethics were utilized and had to be tailored to suit the project context based on discussion within the project (expert hotline; notable focus on ethics in biweekly meetings). This

highlights the importance of a suitable framework, as it saves resources by providing a clear(er) way of approaching ethics in the project context. Otherwise this requires internal effort.

In terms of the activities related to implementing ethics, ethics seemed to ultimately become a part of various project activities, blending in with other project activities, as opposed to being a tacked-on extra responsibility. However, some novel activities remained, such as the expert hotline with an AI expert, which would translate into ethics maintenance costs going forward. In addition, we noted that the implementation of ethics resulted in extra project documentation related to ethics. In part, this extra documentation was a result of ethics being a foreign topic for most stakeholders and necessitated in-depth explanation within the project.

4.3 Case 2

Case 2 summary:

- **Project context:** Data from a proof-of-concept ML project in a software company. Predicting tool for the educational domain. Project customer was interested in exploring potential ethical issues in the project.
- **Who implemented ethics:** Entire development team (4). The development team discussed matters with an external ethics expert on a weekly basis. Attendance in these meetings varied from 1 developer to the entire team.
- **Ethical framework used:** The ECCOLA method for implementing AI ethics [16].

Project activities related to ethics (time spent) [stakeholders involved] in case 2:

- Decision to implement ethics made in a design meeting (1h). [Project management, developers, ethics expert]
- Training workshop on using the ethical framework (ECCOLA) (1,5h). [3 ethics experts, entire development team, and 6 potential end-users]
- Ethics kickoff meeting (2,5h). [Ethics expert and entire development team]
- Use of ethical tool during development (?h) [Developers]
- Weekly check-up meetings with ethics expert (1h) [Ethics expert and 1 to 4 development team members]
- Additional ethics documentation and end reporting (1–2 sheets per iteration) [1–4 developers]

Case 2 Observations. Compared to case 1, the decision to implement ethics in case 2 proceeded in a more straightforward manner. As the project was an ML project, it was possible to utilize a method for AI ethics (ECCOLA [16]). This made it easier for the stakeholders to approach ethics in the project context in various ways. I.e., *what* is going to be done and *how*. Consequently, early on in the project, actions related to ethics could be defined more accurately.

However, this did not result in ethics taking notably less resources. In fact, ethics seemed to take up *more* resources, compared to case 1, especially because the implementation of ethics involved more stakeholders in case 2.

Following the larger initial **ethics investment**, the implementation of ethics then proceeded more systematically. Whereas in case 1 the discussion on ethics continued throughout the project between the developer and the ethics expert, in case 2 the implementation proceeded as planned initially.

Going into **ethics maintenance**, the recurring, distinct ethics-related activities were the weekly check-up meetings with the ethics expert. However, as the project progressed, these focused more on the reporting progress rather than guiding discussion. The additional ethics documentation and reporting also continued, although this was not out of necessity, but because the company itself was curious how ethics was being handled in the project. Otherwise, ethics had become a part of the normal development activities of the company.

4.4 Case 3

Case 3 summary:

- **Project context:** Data from project where a design agency (client company) commissioned software from a consultant company. Project customer specifically requested ethical software.
- **Who implemented ethics:** 3 developers and product manager (senior dev.); 4 developers in total.
- **Ethical framework used:** The ECCOLA method for implementing AI ethics [16].

Project activities related to ethics (time spent) [stakeholders involved] in case 3:

- Decision to implement ethics made in a design meeting (1h). [3 client company representatives and ethics expert]
- ECCOLA tutorial, initial training for the used ethics framework (1,5h). [3 ethics experts, entire development team, and 5 customer representatives]
- Ethics kickoff (2,5h). [Ethics expert, entire development team, and 3 client company representative]
- Use of ethical tool during development (?h) [1–4 developers; varied by iteration]
- Weekly project meeting. Ethics was handled like any other requirement in the backlog (1h) [Entire development team and client company 2–5 representatives]

Case 3 Observations. Case 3 followed a similar pattern as the other cases in terms of the initial **ethics investment**. A notable initial investment was required to define what to implement. As the project then progressed, ethics, like in case 2, was incorporated into existing practices (e.g., discussing ethics in weekly project meetings as opposed to separate ethics-related meetings).

However, as the project began to draw to a close, resource optimization was carried out, and as a result, specifically ethics-related activities were cut. This seems to imply that ethics was nonetheless not completely embedded into any

existing processes and some **ethics maintenance** costs remained that warranted cutting. The customer, who had initially requested ethical software, ultimately considered it a secondary priority. It would, thus, seem that the potential **ethics revenues** were not considered worth the resources at this stage of the project.

5 Discussion

This paper furthers the AI ethics body of knowledge through empirical insights. As the field is lacking in empirical studies [12,13], our understanding of *how* AI ethics is implemented in practice is also lacking, which is considered to be a key issue in the area [9]. Through the practical insights from the three cases, we provide an initial look at the practice of AI ethics from the novel point of view of resources and costs, furthering this understanding. By providing an initial look at the cost of ethics in SE, we hope to motivate further interest on the practical questions of AI ethics.

To begin understanding the cost of ethics in SE, and AI ethics specifically, we turned to a software quality cost estimation framework [11], which we tailored for the context of ethics (Sect. 3). In this initial study, we approached the phenomenon through the project activities undertaken to implement ethics, in order to understand what requires resources when implementing ethics. While the framework provided a basis for this initial discussion, more detailed cost estimation frameworks specifically designed for the purpose of (AI) ethics could be developed going forward, if cost estimation becomes an active concern in ethics in SE.

Further on the note of our comparison to quality, akin to the past software quality experts, the implementation of ethics in SE, at this stage, seems to require an investment in ethics experts, external or internal. In all our cases, ethics experts were present throughout the project and actively leveraged for their expertise by the project staff. Canca [2] also argues that an ethics expert is required in the process so that developers can contact them when faced with challenging ethical issues (in this case, 'challenging' as defined by the tool they are proposing). A similar process was seen in our example cases, and especially case 1. Ethics experts, in this case external ones, were included in the project and provided assistance as needed. It would, thus, seem that ethics indeed requires a continuous investment (ethics maintenance).

5.1 Practical Implications

Ethics takes effort (resources). Ethics is still new in SE, and especially the ethical discussion on AI has made ethics a common topic of discussion recently. Implementing ethics into practice is still challenging and established practices and processes are lacking, making resource estimation difficult. This paper provides an initial look at what implementing ethics could mean in practice as far

as project resources are considered, highlighting that ethics requires resource commitment, with a focus on the initial investment.

However, as the practical implementation of (AI) ethics is still an emerging area of research and practice, the practices and processes required to do so may vary greatly between organizations and project contexts. In this regard, we would recommend the use of an ethical framework to guide the implementation of ethics. This can be a set of guidelines or a method, or any other suitable artefact that helps you define *what* is ethics in your project context. If no suitable framework exists for your application context, either use a more generic one (e.g., business ethics) or consider developing one yourself. By having a shared understanding of what ethics means for your project, you can start planning *how* to develop an ethical system.

Values will get implemented in a service whether it is done systematically or not. By actively looking to tackle AI ethics, it is possible to make a conscious, informed decision on which values to implement. Through nonconformance, it is left up to the developers and other stakeholders working on the system to implement their own values as they see fit, consciously or subconsciously.

5.2 Limitations

As these cases were proof-of-concept projects, we are not able to provide insights into *ethics revenues* and only some initial ones into *ethics maintenance* based on this data. Though our data from the three cases was collected over time, a more systematic, longitudinal approach would be required for a more comprehensive study looking at all three types of costs (ethics investment, ethics maintenance, and ethics revenues). In this regard, we also highlight that these are the results of our limited observation access; it is possible that the cases included more activities related to ethics we were not able to document. Nonetheless, given the novelty of the phenomenon, we feel that this paper provides a starting point for investigating AI ethics from a new point of view that is especially of interest to companies looking into AI ethics.

The use of an ethical framework, we argue, is pivotal in implementing ethics in practice in SE, also from a resource estimation point of view. A framework, such as a set of guidelines or a method, helps us define *what* ethics is in the given project context, giving us clear boundaries within which to work. Otherwise, notable effort is spent on defining the relevant concepts before starting, although such work may be required when operating in novel application areas. However, such frameworks arguably impact what is being done to implement ethics or how ethics is implemented as well. Our findings only serve to provide initial insights into what types of activities and resources *may* be needed when implementing ethics, but given the emergent nature of the area, these may vary greatly by project, based on the ethical framework being utilized, among other factors. E.g., guidelines may only contain sets of principles but little practical guidance, while a method might provide a process to utilize.

Finally, this paper simply provides an initial look at the phenomenon. The data we have utilized was not originally collected to evaluate the implementation

of (AI) ethics from a resource point of view, but to develop the ECCOLA method [16]. While we feel that it nonetheless serves as a starting point for studying this phenomenon, it is hardly a comprehensive look at the process of implementing ethics from a resource and business point of view. Some of the projects may have included activities related to the implementation of ethics that we were not able to document based on our data. On the other hand, as we were not explicitly investigating the resource point of view through our observation and other data collection, it could be argued to not have biased the results by motivating a more extensive investment. Ultimately, the goal of this data was simply to demonstrate the otherwise conceptual points of this paper.

6 Conclusions

In this paper, we provided initial insights into the cost of AI ethics. The current state of AI ethics is reminiscent of how software quality was approached in the early 2000s. Often overlooked at the time, quality still had long-term consequences for software, was costly to implement in production, and was an interdisciplinary endeavor involving various stakeholders, much like AI ethics today. Over the decades, quality evolved from simple quality assurance to a continuous process embedded into organizational culture. Only time will tell whether AI ethics will also mature in the same way.

We adapted a framework for software quality cost estimation into the context of (AI) ethics after drawing parallels between AI ethics and software quality to justify doing so. Based on the framework, we proposed a similar cost framework for the implementation of AI ethics. We then utilized empirical data from three cases to elaborate on the proposed framework by providing an initial look at what types of activities result in the associated costs. Based on the empirical examples, ethics in SE seems to require a notable initial *ethics investment* (e.g., initial training and planning), followed by *ethics maintenance* (e.g., due to the continued involvement of ethics experts). However, the project activities related to ethics may vary between projects, and especially depending on the ethical framework used to guide the process, as tools such as methods may propose specific practices in SE, while tools such as ethical guidelines may necessitate internal effort to devise relevant processes and practices.

As for future research, the practical implementation of AI ethics remains a challenge. Overall, we urge further empirical studies into AI ethics in general, especially ones focusing on practices, methods, and processes for bringing AI ethics into practice. While we certainly urge further studies into the cost of AI ethics as well, for which this paper lays some initial groundwork for, we feel that a better understanding of *how* AI ethics is implemented is also required in this regard. It is arguably far easier to conduct resource estimation for a clear process than it is to do so for ad hoc implementation of AI ethics.

Acknowledgments. This work was partly funded by local authorities (“Business Finland”) under grant agreement ITEA-2020-20219-IML4E of the ITEA4 programme.

References

1. Barraquier, A.: Ethical behaviour in practice: decision outcomes and strategic implications. *Br. J. Manag.* **22**, S28–S46 (2011)
2. Canca, C.: Operationalizing AI ethics principles. *Commun. ACM* **63**(12), 18–21 (2020)
3. Ethics guidelines for trustworthy AI (2019). <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
4. Friedman, B., Kahn, P.H., Borning, A., Huldtgren, A.: Value sensitive design and information systems. In: *Early engagement and new Technologies: Opening up the Laboratory*, pp. 55–95 (2013)
5. Giray, G.: A software engineering perspective on engineering machine learning systems: state of the art and challenges. *J. Syst. Softw.* **180**, 111031 (2021)
6. Gotterbarn, D.W., et al.: ACM code of ethics and professional conduct (2018). <https://www.acm.org/code-of-ethics>
7. Van den Hoven, J.: Moral methodology and information technology. In: *The Handbook of Information and Computer Ethics*, pp. 49–67 (2008)
8. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nat. Mach. Intell.* **1**(9), 389–399 (2019)
9. Mittelstadt, B.: Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* **1**(11), 1–7 (2019)
10. Morley, J., Floridi, L., Kinsey, L., Elhalal, A.: From what to how: an initial review of publicly available ai ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* **26**(4), 2141–2168 (2020)
11. Slaughter, S.A., Harter, D.E., Krishnan, M.S.: Evaluating the cost of software quality. *Commun. ACM* **41**(8), 67–73 (1998)
12. Sloane, M., Zakrzewski, J.: German AI start-ups and “AI ethics”: using a social practice lens for assessing and implementing socio-technical innovation. In: *2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2022*, pp. 935–947. Association for Computing Machinery, New York, NY, USA (2022)
13. Vakkuri, V., Kemell, K., Kultanen, J., Abrahamsson, P.: The current state of industrial practice in artificial intelligence ethics. *IEEE Softw.* **37**(4), 50–57 (2020)
14. Vakkuri, V., et al.: Time for AI (ethics) maturity model is now. *arXiv preprint arXiv:2101.12701* (2021)
15. Vakkuri, V., Kemell, K.-K., Jantunen, M., Abrahamsson, P.: This is just a prototype: how ethics are ignored in software startup-like environments. In: *Stray, V., Hoda, R., Paasivaara, M., Kruchten, P. (eds.) XP 2020. LNBIP, vol. 383*, pp. 195–210. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49392-9_13
16. Vakkuri, V., Kemell, K.K., Jantunen, M., Halme, E., Abrahamsson, P.: ECCOLA-a method for implementing ethically aligned AI systems. *J. Syst. Softw.* **182**, 111067 (2021)
17. Vakkuri, V., Kemell, K., Kultanen, J., Siponen, M.T., Abrahamsson, P.: Ethically aligned design of autonomous systems: Industry viewpoint and an empirical study. *arXiv preprint arXiv:1906.07946* (2019)
18. Watson, T.: Reputation and ethical behaviour in a crisis: predicting survival. *J. Commun. Manag.* **11**(4), 371–384 (2007)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Software Startups



Benefits, Challenges, and Implications of Open-Source Software for Health-Tech Startups: An Empirical Study

Noman Ahmad^(✉) and Nirnaya Tripathi

M3S Research Unit, University of Oulu, Oulu, Finland
{noman.ahmad, nirnaya.tripathi}@oulu.fi

Abstract. Health-tech startups are essential, as they provide cutting-edge solutions to numerous healthcare concerns in the rapidly evolving healthcare industry. They use various technologies to create solutions that boost and advance healthcare systems and healthcare delivery. Open-source software (OSS) technology has become an essential component of startups' toolkits, providing various advantages, such as free access to source codes and opportunities for innovation. Research on OSS in healthcare startups is limited, so our study aims to investigate how health-tech startups perceive the influence of OSS on product development and to identify the challenges they face. To meet this objective, we conducted an empirical study with six health-tech startups, using semi-structured interviews. Thematic analysis was performed on the collected data to identify common themes and subthemes related to the research objective. The findings showed that health-tech startups benefit from the cost efficiency, scalability, and customization of OSS. Open-source software tools, reshape development and promote efficient code management, provide community support, and reduce costs. However, they demand OSS knowledge, management of updates, regulatory compliance, and heightened cybersecurity. Our study adds to the body of knowledge on OSS and healthcare startups and the connection between them. We provide recommendations for health-tech startups, such as embracing OSS tools for their benefits, investing in education and training, and engaging with the OSS community for comprehensive support in their product development processes.

Keywords: startups · health-tech startups · open-source software · product development · empirical study · medical startups

1 Introduction

In today's dynamic digital era, startups and technological entities have been at the forefront of innovation and transformative change. Software Startups focus on crafting software tailored for various sectors, such as finance and education, providing everything from mobile applications to comprehensive enterprise platforms [20]. In the health-tech sector, health-tech startups are revolutionizing healthcare paradigms by leveraging cutting-edge technologies. They utilize different technologies in their product and service

offerings to revolutionize healthcare and develop personalized health strategies [1]. However, while empowering patients, they face competition and inherent challenges, such as the need for more resources. A comparison between software startups and health-tech startups is shown in Fig. 1.







 <p>INDUSTRY TARGET</p> <ul style="list-style-type: none"> • Health-Tech Startup: Focus on the Healthcare sector, offering solutions to improve patient care or healthcare management. • Software Startup: aims to develop software for various industries like finance or education. 	 <p>MARKET DOMAIN</p> <ul style="list-style-type: none"> • Health-Tech Startup: The market is specific to healthcare. • Software Startup: The market can be vast.
 <p>PRODUCT/SERVICE</p> <ul style="list-style-type: none"> • Health-Tech Startup: Offerings may include electronic health records or wearable health tech. • Software Startup: Products can range from mobile apps to enterprise software. 	 <p>REGULATIONS</p> <ul style="list-style-type: none"> • Health-Tech Startup: Faces strict regulations like patient safety or data privacy. • Software Startup: The regulatory might be relaxed.
 <p>STAKEHOLDERS</p> <ul style="list-style-type: none"> • Health-Tech Startup: Involves patients, healthcare providers, or hospitals. • Software Startup: Involves end-users, business analyst, developers, and investors. 	 <p>PRODUCT DEVELOPMENT</p> <ul style="list-style-type: none"> • Health-Tech Startup: Requires rigorous development, which may involve clinical trials. • Software Startup: Development is essential, not very strict.

Fig. 1. A comparison between health-tech startups and software startups

Central to this narrative is the rise and evolution of open-source software (OSS). From its early inception in the 1980s to its widespread adoption today, OSS has profoundly altered the product development landscape by promoting reusability, enabling free access to software source codes, encouraging collaborative contributions, and granting unparalleled freedom to its users [8]. Open-source software offers numerous benefits, including cost savings, enhanced security, and customization [10].

Health-tech products and services have gained importance because of their potential to enhance healthcare infrastructure. Integrating technology with healthcare solutions can improve care quality, foster innovative systems, and reduce costs [14]. In this domain, OSS can aid in areas such as electronic health record (EHR) systems and clinical decision support. Previous studies, such as that by Karopka et al. [11], have highlighted the advantages of OSS in healthcare, citing cost savings, flexibility, and improved interoperability. Syzykova et al. [18] also emphasized the benefits of open-source EHR systems, emphasizing their role in enhancing patient care and achieving cost savings. Given the growing significance of healthcare, health tech startups can leverage OSS to meet healthcare demands.

Research Problem and Objective. However, despite the importance of OSS and health-tech startups, we found very limited, if any, empirical research on OSS adoption in health-tech startups [11, 21]. For example, the authors in [20] discussed various topics on startups but failed to acknowledge OSS research in the startup context. Similarly, a recent literature review [21] lacked OSS research within health-tech startups. To address this gap in the literature, we carried out an empirical study of the benefits of adopting OSS for health-tech startups and the challenges they encounter during its adoption. To understand the topic, we conducted a background literature search on OSS and health-tech startups (Sect. 2). The study framed three research questions (RQs) to explore the issue, employed a qualitative approach, conducted semi-structured interviews with stakeholders in health-tech startups, and performed a thematic data analysis

(Sect. 3). The findings shed light on the benefits of OSS in enhancing product development and the challenges faced during its adoption (Sect. 4). The study discussed the RQs, provided added value to the literature, offered recommendations for practitioners and suggestions for further research topics, and presented the conclusion (Sects. 5 and 6).

2 Background Literature

2.1 Health-Tech Startups

A startup is described as a “*brand-new business with a cutting-edge technological and innovative business plan*” [12]. Startup entities possess the capability for rapid growth and the potential to scale. Ehsan [6] provided a refined definition of startups, emphasizing innovation, growth potential, and risk embracement. A significant factor distinguishing startups from other firms is their focus on product innovation.

The domain of health-tech startups has seen a surge in activity lately. Typically, these startups are characterized and driven by technological breakthroughs, enhanced healthcare offerings, and an increased drive to achieve premium health outcomes at reduced costs [21].

Startups harness emerging technologies, such as artificial intelligence (AI), machine learning, and telemedicine, to devise novel solutions and transform conventional health-care paradigms [17]. Research indicates that one of the primary strengths of health-tech entities is their ability to employ data analytics to craft tailored, data-informed health solutions [19]. Beaulieu et al. [1] highlighted the competitive landscape for these startups, noting that they not only compete with large established corporations but occasionally utilize the services provided by these industries.

2.2 Open-Source Software and Product Development

The origins of OSS can be traced back to the late 1990s, although the concept of free software had its roots in the 1980s. Perceptions of it have shifted over the decades, transitioning from a niche perspective to a mainstream approach accepted by individuals and firms. [8] As Karopka et al. [11] outlined, OSS empowers users with the freedom to utilize, modify, and disseminate software while granting access to its source codes. In today’s digital landscape, many examples of OSS, such as Android OS, Linux, and Apache, are widely adopted [11]. The current ubiquity of OSS means that several firms now design software by integrating OSS components. The OSS model is collaborative, with creators and users actively contributing to its evolution. However, licensing decisions remain the original developers’ preferences [10].

Spender et al. [16] delved into the determinants driving OSS adoption, emphasizing security, software quality, user experience, costs, effort, societal influences, and operational efficiency. Butler et al. [2] further pinpointed organizational strategies in OSS

evaluation; larger entities often rely on structured frameworks or guidelines, whereas smaller outfits typically leverage collective decision-making steered by their leadership.

OSS has profoundly transformed the product development landscape. Academic inquiries have affirmed that OSS can strengthen software quality, accelerate its production, and promote collective contributions from developers [7].

For instance, Fitzgerald [7] observed that OSS initiatives generally exceed proprietary software in code quality and error minimization, an issue attributed to the extensive community of experts monitoring and refining the code. Nonetheless, OSS integration is full of challenges. Issues involving effective project oversight, intellectual property considerations, and security concerns demand attention [5]. Scacchi et al. [14] emphasized that adopting free OSS in crafting extensive software systems is gaining traction as a viable alternative strategy. This approach shows unique examples of project success, deviating from traditional software development practices, and introduces novel methodologies and paradigms in software creation [14].

2.3 Health-Tech Sector and Open-Source Software

The adoption of OSS within the healthcare sector is accelerating. The OSS development model has been influential because it grants the developer community access to freely available source codes, thus fostering collective contributions [11].

Within healthcare, enterprises leverage OSS to deliver enhanced patient care, foster innovation, reduce costs, and add value to the healthcare framework [11].

However, Butler et al. [2] noted that organizations encounter challenges when integrating OSS components. They need help in crafting efficient operational procedures to evaluate OSS elements. This encompasses estimating the financial implications and risks of adoption, along with concerns about functional requirements and attaching to licensing terms. Given the rapid pace and expansive scale of software development in specific organizations, there is a persistent need to refine software evaluation techniques. While some firms rely on developer-driven strategies and unconventional approaches, others have established systematic protocols to evaluate OSS components, allowing for more detailed and layered assessments.

2.4 Health-Tech Startups, Open-Source Software, and the Research Gap

Based on our review and the available literature [11, 21], there is a need for empirical studies that specifically evaluate the use of OSS in health-tech startups. For instance, a paper by a software startup research network titled “Software Startups – A Research Agenda” [20] acknowledged the omission of OSS as a research topic, which is a limitation of their study. Additionally, a recent literature review of health-tech startups in healthcare service delivery [21] emphasized the transformative impact of technology on healthcare, highlighting quicker treatments, enhanced emergency care, and innovations,

such as telemedicine and e-health. However, the review did not report and address the description of OSS research in the health-tech startup literature. Thus, current research regarding the application of OSS in health-tech startups is very limited and needs to be empirically investigated further. To address this research gap, we conducted an empirical investigation that guided health-tech startups on the advantages of OSS adoption.

3 Research Methodology

In our study, we focused on health-tech startups located in a Oulu city in Finland. Understanding the impact of using OSS in these startups is crucial. This study aims to determine the influence of open-source technological components on health-tech startups, and the challenges that these startups encounter when adopting OSS solutions. We have outlined the RQs in Table 1 to address this goal.

Table 1. Research Questions

RQ	Research Question	Rationale
RQ1	What are the perceived benefits of open-source software for health-tech startups?	With RQ1, we seek to understand the specific advantages or positive aspects of using OSS for health-tech startups
RQ2	In what ways can open-source software improve the product development processes for health-tech startups?	Through RQ2, we aim to explore the practical and strategic implications of leveraging OSS in the product development life cycle of health-tech startups
RQ3	What challenges do health-tech startups face when using open-source software in their product development?	Using RQ3, we intend to identify the potential pitfalls or obstacles that health-tech startups might encounter when using OSS in their product development

3.1 Research Approach

We adopted an empirical research approach using semi-structured interviews to delve into the experiences and viewpoints of interviewees concerning the adoption of OSS technology within health-tech startups. Qualitative research is useful for exploring complex scenarios, such as the incorporation of emerging technologies into organizational settings [4]. The startup's selection criteria depended on their use of OSS technology in product development. Interview participants from healthcare-related startups were selected based on their relevant expertise and background in the domain. We employed purposive and snowball sampling techniques to identify the case companies and select the interview participants. The aim was to identify OSS technology adoption among startups focusing on healthcare solutions. The interviewees included the startups' chief executive officers, product managers, and key decision-makers familiar with integrating and utilizing open-source technology.

3.2 Data Collection

Semi-structured interviews served as the primary means of data collection. Interviews were used because of their adaptability, allowing for a tailored approach to collecting information and resulting in comprehensive and in-depth data [9]. To meet our research objectives, we designed a mix of open- and closed-ended questions to gather insights into the participants' experiences and views on using open-source technology within health-tech startups. The set of interview questions was segmented into three sections. The initial section consisted of introductory questions, collecting information about the participants and their respective startups. The core segment of the interview revolved around questions related directly to our research aims. Finally, the concluding section comprised wrap-up questions. As the discussions progressed, some questions evolved naturally, such as how OSS was integrated into existing systems and its advantages.

For a thorough analysis, each interview was audio-recorded and later transcribed. The participants' consent was obtained for these recordings, and a summary of our findings was shared with them for their approval. Data were collected from six practitioners representing six different health-tech startups. All interviews were conducted via Microsoft Teams, with each interview lasting approximately 45 min. The participants had relevant experience in utilizing OSS in health tech startups. In Table 2, further details are available; for example, startups are denoted by "C" as ID. Furthermore, their business domain, such as Business-to-business (B2B) or Business-to-consumer (B2C), is highlighted. Similarly, their founding year and the number of the startup's employees are mentioned. Finally, the Interviewee ID is denoted with "P" along with their role, and information on the startup's product or service description is stated.

Table 2. Overview of the Health-tech Startups' Characteristics and the Interviewees' Roles involved in the study

Startup ID	Business Model	Year	Startup Size	Interviewee ID	Role	Product/Service Description
C1	B2B	2015	1–10	P1	CEO	Preventive care system focusing on oral health
C2	B2B	2004	1–50	P2	CTO	Patient electronic health records system
C3	B2B/ B2C	2014	1–20	P3	Product Manager	Warehouse and logistics management for healthcare
C4	B2B	2004	1–50	P4	Product Manager	Handheld fundus camera for the retina

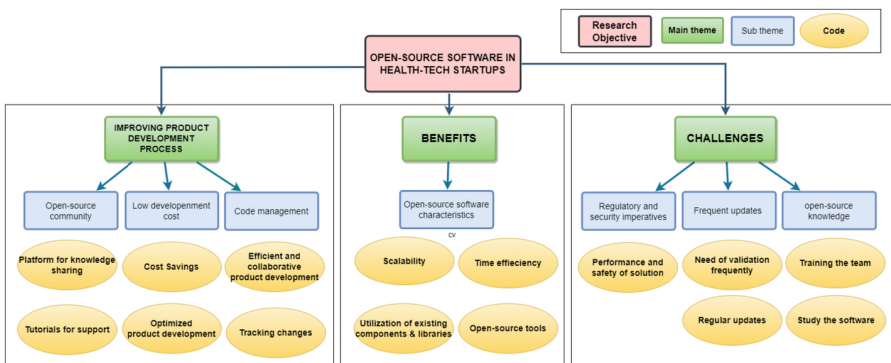
(continued)

Table 2. (continued)

Startup ID	Business Model	Year	Startup Size	Interviewee ID	Role	Product/Service Description
C5	B2B	2015	1–20	P5	CEO	Product related to neurological rehabilitation for speech therapy
C6	B2B/ B2C	2017	1–10	P6	CTO	Patient management systems and healthcare internet of things products

3.3 Data Analysis

Thematic analysis was used to identify, examine, and establish recurring patterns within the data [3]. A systematic approach was taken with the interview transcripts to detect patterns, central themes, and essential insights. The data were organized and categorized using specific codes. Segments of text that represented similar ideas or notions were labeled with these codes. Upon further analysis of the coded data, common themes emerged. Each identified theme emphasized a principal aspect of the research, such as enhancing product development via OSS or the challenges faced when adopting OSS in health-tech startups (see Fig. 2 for code, sub theme, and themes that emerged after data analysis).

**Fig. 2.** Data analysis and thematic results

3.4 Study validity discussion

To ensure the credibility and trustworthiness of our research findings, we used three assessment criteria. These were construct validity, which helped us to measure our study's objective accurately; external validity, which examined the applicability of our findings in real-world settings; and reliability, which aimed to ensure that our research methods and analysis were consistent and dependable. In the following section, we will discuss these three criteria in detail.

Construct Validity. In this study, we developed interview questions aligned with the RQs to ensure construct validity. Additionally, data were gathered from six semi-structured interviews with individuals experienced in health-tech startups and product development. The potential for data inaccuracies because of the interviewer's influence was reduced by conducting numerous interviews. As a result, this research mitigated some potential construct validity risks.

External Validity This research discusses the utilization of OSS in health-tech startups. By incorporating interviews from various health-tech startups, the study minimizes potential biases that might have emerged if it were based on a single interview or company. The sample size was limited to six, and all the startups were based in Oulu, Finland. Therefore, results are confined in their generalizability.

Reliability To ensure reliability, this empirical research provides a detailed explanation of the methodology, data collection, and analysis approach that was used to answer the research questions. However, it's important to note that different researchers may arrive at different outcomes, as the data obtained through semi-structured interviews can be influenced by various factors, including the context and the interviewee's level of knowledge at the time of the interview.

4 Result

We report the insights derived from the data analysis in this section, addressing the study's objectives and answering the RQs.

4.1 Benefits of Adopting Open-Source Software for Health-Tech Startups

Time Efficiency: A recurring theme among the participants was the time-saving advantage of OSS. P5 emphasized that without OSS, they would have had to "start from scratch," which was a time-consuming endeavor. Similarly, P1 highlighted the "faster time to market" benefit, suggesting that their startup, C1, could swiftly introduce their products by leveraging pre-existing OSS. This approach allowed them to focus on innovating unique features rather than reinventing typical ones, which is a benefit particularly useful for health-tech startups with constrained resources.

Scalability: P1 pointed out the scalability inherent in OSS. Scalability ensures software adaptability to fluctuating demands and allows for modifications to specific needs. This adaptability was confirmed by P4, who mentioned building proprietary software on top of OSS and showcasing the scalability potential of OSS.

Utilization of Existing Components and Libraries: Both P1 and P2 emphasized leveraging existing OSS components. By “using existing components instead of writing our code,” as P2 noted, health-tech startups can expedite their development processes. This view does not mean “reinventing the wheel” but capitalizing on the collective efforts of the OSS community. P4 and P5 provided insights into the diverse applications of OSS. For example, C4’s products are embedded in Linux and utilize various OSS libraries. By contrast, C5 focuses on virtual reality simulations, leveraging OSS components from the gaming industry, particularly Unity. These narratives highlight the versatility of OSS across various domains within health-tech startups.

Prominent Open-source Tools: All interviewees highlighted the significance of OSS tools, with a recurrent emphasis on Linux, GIT, Angular, and Android Studio. For example, P1 said that Angular 2 +, Ionic, and Google Technologies underscore the growing trend of using open-source frameworks for mobile and web applications. Diving deeper, P3 elaborated on the multifaceted role of open-source tools, such as the pivotal role of GIT in version control in C3. Similarly, in C4, they used Yocto, a Linux-based tool, and Jenkins, a Java-based DevOps platform, to support the development of the diverse functionalities of their products. In C5, Unity further showcases the expansive open-source ecosystem available to startups, with its community being a valuable resource. By leveraging OSS tools, startups can optimize their development processes, support team collaboration, and properly allocate resources, thus achieving a more streamlined product development and delivery course.

In conclusion, the participants’ descriptions confirmed the pivotal role of OSS in within health-tech startups. The benefits, from time and cost efficiency to scalability, flexibility, and the ability to leverage existing solutions, empower health-tech startups to optimize resource allocation and accelerate development.

4.2 Ways in Which Open-Source Software Improve the Product Development

Most participants discussed the fast pace of product development because of support from the open-source community, as well as time savings because of proper version management of the product. They also mentioned cost reduction, which directly improved product development. Three principal subthemes were identified regarding the impact of OSS technology on product development: support from the open-source community, low development costs, and version management.

Open-Source Community: The research participants frequently mentioned the support they received from the open-source community. P1 emphasized the vast resources available, including tutorials, which offer flexibility in using and modifying OSS. This view stresses the community’s role in aiding developers through valuable insights and

resources. Comprising passionate software enthusiasts, the open-source community provides extensive help, often through experienced developers who share their expertise. P5 highlighted the community's role in offering pre-existing tools, helping save time for developers. The respondents specifically mentioned Unity software's open-source community, which aids game development, and how it played a pivotal role in the creation of C5's product for evaluating attention deficit hyperactivity disorder symptoms through virtual reality simulation. The responses of P1 and P5 regarding the key role of the open-source community in enhancing product efficiency were consistent. The community promotes reusability and continuous development by providing a platform for knowledge sharing, collaboration, and innovation. Health-tech startups can leverage these resources to expedite their development processes, thus avoiding redundant efforts. The open-source community acts as a catalyst, pushing health-tech startups forward by providing them with resources and expertise.

Low Development Cost: Most participants in the study emphasized the significant cost savings associated with OSS technology in the product development process. P1 highlighted the cost-effectiveness of OSS as a crucial advantage, especially for startups. Such software is often free or offered at a minimal cost, reducing the financial strain on developers. P2 stressed the absence of licensing costs when deploying OSS solutions, which is especially beneficial for health-tech startups aiming to keep their operational costs low. C3 and C6 were able to focus on saving by avoiding the purchase of expensive proprietary libraries, thus favoring open-source alternatives. The interviewees' collective responses highlight the transformative impact of OSS on startups, particularly in terms of cost savings. The elimination of hefty licensing fees and the ability to customize software to one's specific needs allow health-tech startups to allocate their resources more effectively. This results in financial savings and fosters innovation, scalability, and sustainable growth.

Efficient Code Management: Code management is pivotal in software development; it facilitates collaboration, tracking of changes, and error prevention. The participants emphasized the significant role of OSS in version management, particularly the use of GIT. In C5, GIT is a core tool used for version management; its importance in tracking source code changes is highly valued. The tool aids in understanding the evolution of a product, ensuring regulatory compliance, and maintaining a clear change history. P3, with a programming background, also endorsed GIT, noting its ease of use when handling code and the recent switch of their start-up to this platform because of the positive feedback on it. Using an open-source tool for version management ensures reliability and stability during product development and saves time and effort.

4.3 Challenges in Adopting Open-Source Software

The challenges while adopting the OSS theme include frequent updates, OSS knowledge, and regulatory and security aspects.

Frequent Updates: In health-tech startups, the rapid evolution of OSS presents significant challenges, as P3, P5, and P6 highlighted. They identified regular updates as a

primary concern. While beneficial for software enhancement, these updates can disrupt development and validation processes. P3 emphasized the importance of understanding software to anticipate and manage these updates, noting that such changes can introduce complexities requiring time-consuming modifications. P6 elaborated on the challenges posed by updates, stressing that open-source frameworks often undergo annual revisions. This swift pace complicates the development process, sometimes necessitating a freeze to ensure consistency with the chosen framework version. Beyond development, P6 also highlighted the intricacies of application validation between updates. Ensuring that applications meet functionality, compliance, and performance standards becomes difficult, as each update might introduce changes that demand rigorous testing and verification. Adding to this challenge, P5 mentioned that the frequent updates inherent in open-source technologies result in the need to continuously validate them. Frequent changes can compromise the reliability and stability of applications, especially given the rapid pace of upgrades. While OSS offers numerous advantages, health-tech startups must navigate the inherent challenges that come with them. These include managing consistent updates, pausing development for stability, ensuring rigorous application validation, and enabling swift adaptation to updates. The participants highlighted the need for health-tech startups to be proactive and strategic when integrating OSS into their operations.

Open-Source Software Knowledge: Open-source adoption in health-tech startups presents opportunities and challenges. A recurring theme among the participants was the steep learning curve associated with integrating OSS. P1 highlighted that unfamiliarity with OSS can slow down the development process. This view was supported by P4, who faced challenges in getting their team on board because of a lack of prior experience with open-source tools. Such challenges underscore the need for health-tech startups to invest in training and expertise in order to ensure seamless integration and effective collaboration. Another significant concern is the integration of open-source technologies with existing proprietary systems. As P1 pointed out, mismatches between the two can lead to technical issues, further delaying development. Health-tech startups must understand in depth the software they are integrating and invest in specialized expertise to navigate potential integration hurdles. Vulnerabilities in open-source components are another area of concern. P2 and P5 emphasized the importance of understanding the life cycles of open-source components and being aware of their vulnerabilities. Regular updates, while essential for security and functionality, can be challenging. As P3 noted, frequent updates, although beneficial, can strain resources and complicate the development process. These challenges have added significance for health-tech startups, in which patient data and system reliability are paramount. In essence, while OSS offers cost-effective and flexible solutions, health-tech startups must approach its adoption with caution, preparation, and a commitment to continuous learning.

Regulatory and Cybersecurity Imperatives: Regulatory challenges are pivotal when integrating OSS, especially in sectors such as healthcare. Both P5 and P6 emphasized the significance of security and performance in this context. P5 stated that their startup, C5, constantly evaluated the impact of open-source technology on the safety and performance of solutions. The respondents highlighted the need to determine whether OSS

technologies are integral to the system or merely serve as supplementary tools. This distinction is crucial in deciding compliance with regulatory standards. P6, on the other hand, highlighted the increasing importance of cybersecurity, especially with the proliferation of AI. As AI becomes more embedded in systems, the demand for robust security in OSS intensifies. The open accessibility of such software, while fostering innovation, can also introduce vulnerabilities. The insights from the participants underscored the dual-edged nature of OSS. While it offers flexibility and a vast pool of resources, it also demands rigorous scrutiny, especially in sectors governed by stringent regulations. The integration of AI amplifies security imperatives. It accelerates AI advancements but also necessitates heightened cybersecurity measures.

5 Discussion

In this section, we discuss the RQs, present their added value to the literature, provide recommendations to practitioners, and suggest further research avenues.

5.1 Answers to the Research Questions (RQs)

Open-source software has become an essential component for startups, offering a multitude of advantages that often surpass the difficulties associated with it. Our thorough analysis, based on extensive interviews and data, highlights the vital role of OSS in health-tech startups. While most startups use OSS in a similar manner, they vary in the tools they choose to implement. Table 3 provides a summary of our answers to the RQs.

Table 3. Summary of answers to the RQs

Category	Description	Source
Perceived benefits of OSS for health-tech startups (RQ1)		
Time efficient	With OSS, health-tech startups do not have to build software foundations from the ground up. They can leverage existing frameworks and libraries, thus accelerating their development processes	P1, P2, P5
Scalability	Open-source software offers unparalleled scalability and flexibility. Health-tech startups can easily modify and upscale their operations without significant hurdles	P1, P4

(continued)

Table 3. (continued)

Category	Description	Source
Utilization of existing components and libraries	Unlike proprietary software, OSS allows health-tech startups to tailor the software to their needs, ensuring a more personalized user experience	P1, P2, P4, P5
Prominent OSS tools	Several open-source tools have emerged as game changers for startups. Linux, an open-source operating system, has been widely adopted for its reliability. GIT stands out for code management, whereas frameworks such as Angular cater to web application development	P1, P2, P3, P4, P5, P6
Impact on product development (RQ2)		
Efficient code management	Tools such as GIT streamline version and code management, ensuring the efficient tracking of changes and maintenance of version histories	P3, P5
Open-source community	The open-source community is a goldmine of resources. Health-tech startups can access a wealth of knowledge and expertise from tutorials to forums. This community-driven approach fosters collaboration and continuous learning	P1, P5
Low development cost	One of the primary aspects of OSS in product development is its low development cost. Many open-source tools are free or have minimal costs, providing startups, especially those in nascent stages, with financially viable solutions	P1, P2, P6

Challenges with OSS (RQ3)

(continued)

Table 3. (continued)

Category	Description	Source
Open-source software knowledge	Integrating OSS often comes with a learning curve. Health-tech startups need to acquaint themselves with the distinct structure, workflow, and paradigms of OSS, which can sometimes be starkly different from those of proprietary software	P1, P2, P4
Frequent updates	Open-source software is dynamic, with regular updates and fixes. While these updates introduce new features, they can also pose challenges. Integrating these frequent updates can be resource intensive and time consuming	P2, P5, P6
Regulatory	Health-tech startups must navigate the regulatory landscape, ensuring performance and safety standards compliance. This is especially pertinent when considering the integration of OSS technologies into core systems	P5, P6
Cybersecurity	Cybersecurity has taken center stage with the proliferation of artificial intelligence (AI) technology. The open nature of OSS, while fostering collaboration, can also introduce vulnerabilities. As AI becomes more ubiquitous, ensuring robust security measures becomes paramount	P5, P6

5.2 Theoretical Contributions to the Literature

The findings of our study on incorporating OSS into the growth of healthcare startups align with earlier findings in various crucial aspects. A comparison with prior research reveals notable similarities and insights, which are discussed below.

Karopka et al. [11] and Santarsiero et al. [13], have identified an increasing trend in OSS adoption in healthcare. Our research confirms this, emphasizing the importance of OSS in fostering innovation, reducing costs, and adding value to the healthcare landscape

within health-tech startups. Karopka et al. [11] highlighted the flexibility of OSS, granting users the freedom to access, distribute, and modify its content, especially its source codes. Our findings expand this by illustrating that health-tech startups derive substantial advantages from the transparent nature of OSS and its associated tools. Interestingly, our study introduces new perspectives, such as the role of OSS in product development and the tools that assist startups in managing code modifications.

Similarly, Shaikh et al. [15] and Butler et al. [2] pointed out the challenges of adopting OSS. Some of these are the same as those identified in our research. Both studies drew attention to difficulties, such as the pronounced initial learning phase, the unfamiliarity of OSS, navigation of constant updates, compliance with established protocols, and security concerns. Our findings underscore the need for health-tech startups to recognize the potential risks of OSS adoption and to conduct thorough evaluation and planning before its introduction. One particular challenge that has not been extensively covered in earlier works pertains to the depth of understanding required for OSS. This often necessitates health-tech startups investing in training on OSS, which demands time and resources.

5.3 Recommendations for Practitioners

Based on the results, we recommend that health-tech startups start adopting OSS to increase the efficiency of their products. They should consider using OSS tools, as these provide affordable options, scalability, flexibility, and time-saving benefits. Health-tech startups can use configurable software and existing infrastructure and make their development processes more efficient by utilizing these technologies.

Training and Education: Health-tech startups should start investing in training and education about OSS for their team members because understanding the architecture, workflow, and paradigms of OSS is essential for successful implementation. Health-tech startups can reduce the learning curve associated with adopting open-source technologies by providing proper training and assistance.

Updates and Integration: Health-tech startups should learn how often OSS updates itself and determine whether they want to integrate the updates into their systems. If OSS is updated rapidly, health-tech startups may encounter difficulties adapting their systems to the changes.

Risk Assessment: Health-tech startups should also carefully consider and adhere to any regulatory obligations on using OSS and considering performance, safety rules, and security procedures. They should carry out a thorough risk analysis before deploying OSS. This entails knowledge about the vulnerabilities and difficulties linked to open-source technologies.

Community Engagement: Health-tech startups should actively interact with the open-source community for advice and support. The enormous open-source community makes many resources, courses, forums, and professional opinions available. Health-tech startups may overcome obstacles, learn best practices, and accelerate their growth by utilizing the expertise and experiences of the community.

5.4 Study Limitations and Future Research

The study focused on health-tech startups, yielding a limited sample size of just six startups. This small size affects the broader applicability of the results, although they align with previous research. While the findings offer essential insights, they capture only some aspects of startups' open-source adoption. The potential effects on creativity, teamwork, and competitive advantage require further exploration. In future studies, health-tech startups' experiences with using proprietary solutions could provide a more profound understanding of the unique advantages and challenges of OSS. Additionally, an in-depth look into the security measures employed by health-tech startups when using OSS would be beneficial.

6 Conclusion

Health-tech startups have increasingly embraced OSS for its cost and time efficiency, scalability, and customization. Notable OSS tools are revolutionizing the development processes and code management of startups. However, startups also face challenges despite the numerous advantages of OSS, such as understanding OSS dynamics, managing frequent updates, adhering to regulations, and ensuring cybersecurity. Previous studies corroborate these findings, emphasizing the role of OSS in fostering innovation and cost savings. Health-tech startups are advised to invest in training, understand update cycles, assess risks, and engage with the OSS community to maximize the OSS benefits they obtain.

References

1. Beaulieu, M., Lehoux, P.: Emerging health technology firms' strategies and their impact on economic and healthcare system actors: a qualitative study. *J. Innov. Entrepreneurship* **7**, 11 (2018). <https://doi.org/10.1186/s13731-018-0092-5>
2. Butler, S., Gamalielsson, J., Lundell, B., et al.: Considerations and challenges for the adoption of open-source components in software-intensive businesses. *J. Syst. Softw.* **186**, 111152 (2022). <https://doi.org/10.1016/j.jss.2021.111152>
3. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement, pp. 275–284. IEEE, September 2011
4. Creswell, J.W.: *Research Design Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications Inc, Thousand Oaks (2014)
5. Crowston, K., Wei, K., Howison, J., Wiggins, A.: Free/libre open-source software development. *ACM Comput. Surv.* **44**(2), 1–35 (2008). <https://doi.org/10.1145/2089125.2089127>
6. Ehsan, Z.-A.: Defining a startup - a critical analysis. *SSRN Electron. J.* (2021). <https://doi.org/10.2139/ssrn.3823361>
7. Fitzgerald.: The transformation of open source software. *MIS Q.* **30**(3), 587 (2006). <https://doi.org/10.2307/25148740>
8. Gonzalez-Barahona, J.M.: A brief history of free, open-source software and its communities. *Computer* **54**(2), 75–79 (2021). <https://doi.org/10.1109/mc.2020.3041887>

9. Hove, S.E., Anda, B.: Experiences from conducting semi-structured interviews in empirical software engineering research. In: 11th IEEE International Software Metrics Symposium (METRICS 2005) (2005). <https://doi.org/10.1109/metrics.2005.24>
10. Kapitsaki, G.M., Tselikas, N.D., Kyriakou, K.-I.D., Papoutsoglou, M.: Help me with this: a categorization of open-source software problems. *Inf. Softw. Technol.* **152**, 107034 (2022). <https://doi.org/10.1016/j.infsof.2022.107034>
11. Karopka, T., Schmuhl, H., Demski, H.: Free/Libre open source software in health care: a review. *Healthc. Inf. Res* **20**(1), 11 (2014). <https://doi.org/10.4258/hir.2014.20.1.11>
12. Krejčí, M., Strielkowski, W., Čabelková, I.: Factors that influence the success of small and medium enterprises in ICT: a case study from the Czech Republic. *Verslas: Teorija ir Praktika* **16**(3), 304–315 (2015). <https://doi.org/10.3846/btp.2015.521>
13. Santarsiero, F., Schiuma, G., Carlucci, D., Helander, N.: Digital transformation in healthcare organisations: the role of innovation labs. *Technovation* **122**, 102640 (2023). <https://doi.org/10.1016/j.technovation.2022.102640>
14. Scacchi, W., Feller, J., Fitzgerald, B., Hissam, S., Lakhani, K.: Understanding free/open source software development processes. *Softw. Process Improv. Pract.* **11**(2), 95–105 (2006). <https://doi.org/10.1002/spip.255>
15. Shaikh, M., Vayani, A.H., Akram, S., Qamar, N.: Open-source electronic health record systems: a systematic review of most recent advances. *Health Inf. J.* **28**(2), 146045822210998 (2022). <https://doi.org/10.1177/14604582221099828>
16. Spender, J.-C., Corvello, V., Grimaldi, M., Rippa, P.: Startups and open innovation: a review of the literature. *Eur. J. Innov. Manag.* **20**(1), 4–30 (2017). <https://doi.org/10.1108/ejim-12-2015-0131>
17. Stahl, M., Zarco-Jasso, H., Miralles, F.: How can innovation intermediaries enable product innovativeness in startups? A configurational approach. *J. Open Innov. Technol. Market Complex.* **9**(1), 100007 (2023). <https://doi.org/10.1016/j.joitmc.2023.02.004>
18. Syzdykova, A., Malta, A., Zolfo, M., Diro, E., Oliveira, J.L.: Open-source electronic health record systems for low-resource settings: Systematic review. *JMIR Med. Inf.* **5**(4) (2017). <https://doi.org/10.2196/medinform.8131>
19. Tripathi, N., Oivo, M., Liukkunen, K., Markkula, J.: Startup ecosystem effect on minimum viable product development in software startups. *Inf. Softw. Technol.* **114**, 77–91 (2019). <https://doi.org/10.1016/j.infsof.2019.06.008>
20. Unterkalmsteiner, M., Abrahamsson, P., Wang, X., Nguyen-Duc, A., Shah, S.M.A., Bajwa, S.S., Yagüe, A.: Software startups--a research agenda (2023). arXiv preprint [arXiv:2308.12816](https://arxiv.org/abs/2308.12816)
21. Chakraborty, I., Ilavarasan, P.V., Edirippulige, S.: Health-tech startups in healthcare service delivery: a scoping review. *Soc Sci Med* **278**, 113949 (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Corporate Startups: A Systematic Literature Review on Governance and Autonomy

Konstantin Garidis¹ (✉) , Alexander Rossmann² , and Alan Murray¹

¹ University of the West of Scotland, High Street, Paisley PA1 2BE, Scotland, UK
konstantin.garidis@reutlingen-university.de,
alan.murray@uws.ac.uk

² Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany
alexander.rossmann@reutlingen-university.de

Abstract. Many incumbents observe the startup world in jealousy of their agility and innovational performance. An increasing number of initiatives aim to mimic startup-like procedures in order to increase the incumbents' innovational output. Structural models like accelerators, spinoffs, incubators, or corporate venture capitals aim to achieve that goal by implementing different governance setups. However, the success of such initiatives often remains unclear. While there is broad research on such topics, a clear empirical view on governance mechanisms for entrepreneurial structures in incumbents is missing. This paper outlines how to build a governance model based on empirically validated mechanisms and their relationship to corporate startup autonomy. This is achieved by following the systematic literature review approach by Webster and Watson combined with qualitative data analysis techniques. The results describe relevant gaps in current research and identify promising pathways for future research.

Keywords: corporate startup · corporate entrepreneurship · governance · autonomy

1 Introduction

New and disruptive digital business models enter every market. Over the years, the speed of development and market entry has continuously increased. With the development of new ideas, and thanks to the maturing internet technology and the spreading of digital products in most industries, concepts are designed and tested on the market even faster. These methods of rapid development and introduction of disruptive digital business models are mostly said to be done by digital startups and tech firms [4]. As business model innovation is a new way to create, deliver or capture value [32], it also calls for structural, operational, or cultural renewal [31]. Digital startups inhibit this approach in their essence as they are “an organization formed to search for a repeatable and scalable business model” [1]. Therefore, research and practice mainly attribute the ability to drive digital business models to startups, startup-like structures, and big tech firms [2]. As these abilities are intertwined with a firm's organizational structure, many

incumbents realized a need for autonomous startup-like structures to reach the agility, speed, and flexibility needed. Hence, the idea of corporate startups (CS) has risen. Today incumbents apply many CS models following different strategies; e.g., Weiblen and Chesbrough [39] describe engagement models according to the direction of the innovation flow outside-in or inside-out and equity involvement. Most models today aim to build an environment that enables innovation by offering a certain degree of autonomy from the established structures of the incumbent [31]. Debates have arisen on how incumbents can grant autonomy to their CS, while still maintaining a mutually beneficial relationship, as research has shown that incumbents struggle with professionalizing their CS initiatives [33].

Over time, the topic has also been of high research interest. Currently, various studies are analyzing the effects of implementing specific models like accelerators or incubators [12]. Most researchers investigate these models and their circumstances [6]. Some look into the economic aspects of corporate venturing [7], and others analyze the cooperation or collaboration between the uneven partnerships of startups and corporates [9]. Research has addressed the challenge of utilizing resources from the incumbent or enabling knowledge inflow and outflow while allowing the CS to act autonomously and evolve under the debate of the structural autonomy of CSs. However, the results in this research stream are contradictory [5, 10, 19]. Some research shows that structural autonomy is needed to secure fast and independent decision processes [22]. In contrast, other studies show that CS autonomy (CSA) can hinder resource provision and knowledge flow [20]. Moreover, the success of CS initiatives often remains unclear. As Kötting [18] describes, “a major decision with the implementation of corporate incubation is the degree of autonomy.” There seems to be a “tug of war” between granting autonomy and effectively governing CSs. Additionally, most studies focus on autonomy as a single construct rather than complex governance structures. Conclusively, our research thrives on answering the following questions:

1. Which governance aspects of corporate startups exist in empirical research?
2. How can autonomy be managed from a governance perspective?
3. What research is missing to provide incumbents with an effective corporate startup governance framework?

This study uses a literature review approach to identify the current state of research on the governance aspects of CS models based on the typology by Weiblen and Chesbrough [39]. While there are literature reviews on the organizational aspects of CSs, some address specific models like accelerators [6, 23, 35] or do not focus on governance mechanisms [26, 28, 40]. This review shows, that no study investigates CS governance as a whole. This leads to the current body of knowledge where, although we know about aspects of CS models, how firms implement these models by applying governance mechanisms is still unknown. Our review fills this gap by developing a governance model built on empirically identified mechanisms extracted from the literature using qualitative text analysis and the software maxqda. The model developed by this review enables firms and researchers to investigate CS models from a governance perspective and understand how an optimal configuration could look like.

2 Theoretical Foundation

A startup is a temporary organization and the sole purpose of a startup is to develop and test a new business model [1]. As their purpose is to test new concepts, they need to be able to adapt and develop, based on previously gained experiences. They are usually small and relative newcomers to the market. Hence, these firms typically have no established functional structures like human resources, sales channels, or partners [9].

As incumbents recognize the advantages in agility and flexibility that startups have, they aim to combine their strengths to enhance innovation output. Due to their nature, incumbents optimize their structures, processes and operations to optimally execute their current business model [21]. These structures are needed to optimize operational costs and speed up standardized processes. In recent years, incumbents have increased their efforts to build structures that enable digital business model innovation [6].

Research and practice generally refer to these startup-like structures as CS. A CS shares a startup's attributes, but differs in that it is associated with a corporate incumbent by ownership, strategic partnership, or integration into the corporate structure. The concept of the CS tries to benefit from the agility, and change-embracing structure that startups have, combined with the resources and established processes an incumbent has built. The gap that separates the incumbent and the CS varies hugely [37]. Various attributes of the collaboration, such as ownership, integration into the corporate structure, or even the headquarters' location, determine how deeply integrated the CS is into the incumbent. How such structural attributes affect the abilities of the CS has yet to be researched [18].

While there have been studies on the effects of organizational and structural mechanisms of CS on performance, the existing studies show mixed results. Some scholars advocate a more autonomous CS setup [10]. Other empirical research found evidence that more integrated configurations can benefit CS performance [37]. However, it is still not fully understood how various governance mechanisms can be utilized to manage CSA.

2.1 Corporate Startups Defined

Incumbents follow different CS models and strategies to pursue their innovation goals. Over the years, several of these models have become established in practice. A plethora of research exists to describe distinct models and their attributes [23, 35]. Although these concepts are valuable for analyzing the respective CS model, a typology encompassing all models is needed to investigate the applied governance mechanisms. Weiblen and Chesbrough's approach explains different models by classifying CS models following the innovation flow and equity involvement [39]:

Inside-Out models: Corporate Incubation is often nested into a structured program where internal innovation processes are streamlined into a more agile entity. Firms usually apply these models for innovations that differ too much from the core business, hinting at a need for structural autonomy. Startups emerging from this type are often called spinoffs. The term incubation is also used for outside-in entities that cooperate with startups by providing facilities, mentoring, and other services [15].

Outside-In models: Corporate Venturing describes a well-established model of investing in existing startups according to a strategic goal set by the corporate entity. The process involves individual steps like scouting for fitting startups or comprehensive due diligence. A *Startup Program* is a model used to make promising innovations and products by startups available for the offering corporate. The format allows the incumbent to engage with several startups and explore possibilities. In exchange, the startups receive benefits like consulting, or access to the corporate ecosystem.

2.2 Autonomy and Governance

Autonomy has been a topic of debate in CS research for quite a while now. Many studies suggest that a CS needs a certain level of autonomy to enhance its learning and develop innovation capability fully. This idea is substantiated by structural ambidexterity, which suggests separating organizational structures into entities according to the two objectives of exploiting existing markets and exploring new ones [34]. The idea of CSA is to create an environment for the CS that promotes creativity and flexibility to enable exploration [20]. Other research shows that a high degree of autonomy can adversely affect CS performance as it impedes knowledge inflow from the CS to the parent firm [5, 16].

There seems to be a “tug of war” between granting autonomy to create a creative environment that promotes exploration, and setting up structures and processes that integrate the CS into the parent to secure alignment between the two. Researchers have addressed this issue by distinguishing different types of autonomy: Structural autonomy refers to the extent to which a CS is separated from its parent [3]. Operational autonomy describes the extent to which CS operations, such as human resources, are shared with the parent firm [11]. Planning autonomy represents the strategic aspects of autonomy and describes the CS ability to autonomously set its goals and strategic directions [16].

Autonomy is a complex construct influenced by various mechanisms and their interplay [37]. Research has established similar dimensions in governance research as they address comparable design dimensions of a firm: structures, processes and operations, and relational mechanisms [14, 36]. Studies show that effective governance mechanisms can significantly improve a firm’s performance. Although the relationship between CSs and their parent has been studied extensively [27, 30], research is just starting to utilize the mentioned governance dimensions in the context of CSs.

3 Research Approach

We follow a systematic literature review process by Webster and Watson to analyze the body of knowledge on CS [38]. The review aims to identify related work on governance mechanisms and their impact on CSA to understand how an optimal CS governance setup may be designed. The research process follows five phases. Table 1 summarizes the results of the process.

Phase 1 Search: Each selected search string in table 1 represents a CS model based on the conceptual framework described in Sect. 2.1. These search strings ensure that we include studies for all CS models to build a broadly applicable framework. Additionally, we added a general search string to ensure the inclusion of studies on general CS models.

We conducted the title and abstract search and used the mechanisms provided by the databases in Table 1 to ensure that plurals and differences in spelling, e.g., “incubation” vs. “incubator” are included. As CS models are recently gaining more attention, we searched for studies published in peer-reviewed journals and conferences, as the latest research is usually first published at conferences. To ensure that the studies we found truly represent the current phenomena of CSs, we omitted studies published before 2010 from the search. Thus, 883 papers were identified for the next step.

Phase 2 Evaluation: This phase represents the title and abstract review. After removing duplicates, 556 studies remained for further evaluation. Only studies that empirically analyze or develop structures and governance mechanisms of CS and their effect on CSA or its performance effects are selected. We excluded conceptual papers [25] or studies that don’t focus on CS governance mechanisms from the review [24]. At this stage, 58 papers remain for further analysis.

Phase 3 Reading: This phase represents the full-text review. During this process, we excluded some papers due to their lack of focus on governance mechanisms and we found two additional papers through forward-and-backward search. Finally, 12 studies remained for assessment.

Phase 4 Coding: We quantitatively extracted governance mechanisms using the analyzing software maxqda analytics pro. We only coded mechanisms in the results presenting sections, discussion, and conclusion to ensure that the model only includes empirically identified mechanisms from the literature. This restriction ensures that non-empirical ideas or examples do not compromise the final model. The model separates the mechanisms according to the established governance framework we previously described and divides them into the innovation flows if applicable [36].

Phase 5 Writing the Review: We combined the identified mechanisms from the previous phase into our model. All mechanisms found in the last step are mapped to the three dimensions of the governance framework by Vejseli [36]. After completing the model-building, the review describes the knowledge base for each mechanism, and we discuss their implications, effects on CSA and define gaps in the model.

4 Descriptive Results

In the context of framework development, different aspects are essential to address. Table 2 lists the twelve identified studies, their investigated CS model, and innovation flow. To understand how incumbents govern these models, we map the models with the governance mechanisms and autonomy aspects, respectively. Most studies combine governance and autonomy explicitly. The table shows they investigate similar governance and autonomy dimensions, e.g., *structural governance mechanisms* and *structural autonomy* [5, 37]. Some studies incorporate aspects of autonomy implicit as an attribute of the investigated governance mechanisms [26, 29]. This circumstance is especially evident for *structural autonomy* aspects like holding equity or general statements on “structural separation” [29].

While most studies examine *structural autonomy* in their research, all studies investigate *operational governance* aspects. This imbalance might indicate a blind eye in CS governance research on the other dimensions. Seven of twelve articles were published

Table 1. Search process

Innovation Flow	Keyword	Science Direct	WoS	Business Source Ult	Emerald	T&F	Sum
Phase 1: Search							
General	corp. Startup	2	15	14	5	8	44
Inside-out	corp. Spinoff	14	50	35	3	2	104
Inside-out	corp. Incubation	0	103	35	48	13	199
Outside-In	corp. Accelerator	26	71	64	29	8	198
Outside-In	corp. Venture capital	63	130	121	9	15	338
Sum:		105	369	269	94	46	883
Phase 2: Evaluation							
Sum without duplicates		556					
Title and abstract review		58					
Phase 3: Reading							
Full-text review		12					

in the last three years, and only one identified study was published before 2015 [41]. The fact that most studies use qualitative research methods and their recent publication dates indicate that investigating CS through the lens of governance mechanisms and autonomy seems to be a relatively new aspect of CS research. However, researchers in CS research seem to prefer qualitative methods due to data availability issues for quantitative methods [10]. The explorative stage of the research stream strengthens the argument for conducting this literature review to build a holistic governance model.

Although the selection process excluded studies only containing distinct governance mechanisms, just four of the twelve articles investigated all three established governance dimensions. All three studies having all three governance dimensions only implicitly investigate the role of *structural autonomy*, excluding the other autonomy dimensions [17, 19, 22]. The findings show that research has only studied fractions of CSA.

There is no imbalance in the number of studies addressing the two directions of innovation flow. Although there were more search results for the outside-in search terms, as shown in table 1, the resulting papers equally focus on inside-out and outside-in models. The fact that there are more inside-out studies proportionate to the search results could hint that CS governance is more eminent in inside-out research.

5 Corporate Startup Governance Framework

This section represents *phase 5* of the review. The framework presented in Table 3 provides all mechanisms identified in *phase 4*. We sorted the mechanisms based on the number of occurrences and referenced the respective sources for each mechanism. Furthermore, the table maps the respective autonomy dimensions described in the studies, if applicable. In the following, we illustrate the framework by describing the mechanisms for each dimension and how they are related to CSA. Where there is a difference between inside-out and outside-in models, we state it in the description. Section 5.4 describes how the literature defines each autonomy dimension and the interplay between governance mechanisms and CSA.

Table 2. Studies on corporate startup governance and autonomy

Study	Innovation Flow	CS Model	Method	Governance	Autonomy
[10]	Inside-Out	Internal Corporate Venture	Quantitative	Operational	Planning
[26]	Inside-Out	Internal Corporate Accelerator	Qualitative	Structural; Relational; Operational	Structural
[5]	Inside-Out	Internal Corporate Venture	Quantitative	Structural; Operational	Operational
[37]	Inside-Out	Corporate Venturing	Qualitative	Structural; Operational	Structural; Operational; Planning
[19]	Inside-Out; Outside-In	Corporate Incubation	Qualitative	Structural; Relational; Operational	Structural;
[29]	Inside-Out; Outside-In	Corporate Incubation	Qualitative	Structural; Operational	Structural;
[8]	Inside-Out; Outside-In	Corporate Incubator	Quantitative	Relational; Operational	Relational
[41]	Outside-In	Corporate Venture Capital	Quantitative	Structural; Operational	Structural; Operational
[17]	Outside-In	Corporate Accelerator	Qualitative	Structural; Relational; Operational	Structural

(continued)

Table 2. (continued)

Study	Innovation Flow	CS Model	Method	Governance	Autonomy
[22]	Outside-In	Corporate Accelerator	Qualitative	Structural; Relational; Operational	Structural; Relational
[42]	Outside-In	Incubator	Quantitative	Operational	Operational; Relational
[13]	Outside-In	Corporate Accelerator	Qualitative	Relational; Operational	Relational; Planning

5.1 Structures

The *management* dimension describes the degree of support and participation of the incumbent’s management in the CS. The weakest form of management participation is management attention; a situation where the management is not actively involved but aware of the CS. Management attention is the first stage in gaining management sponsorship [17, 37]. All studies agree that strong management sponsorship and commitment represent a vital success factor for CSs [17, 26, 29, 37]. This assessment is different in the case of management influence and involvement. Management influence describes a situation in which the management does not actively participate in the CS, but has the power to influence its strategies and operations. This influence could be beneficial, depending on the management’s knowledge about the CSs operations and market [37]. There are contradicting results in the case of active management involvement. Although Waldkirch et. al. [37] found positive effects in different circumstances, Yang [41] identified adverse effects of active management involvement and CS performance. Strong management backing helps the CSs get the necessary resources and freedom, thus improving their performance. In contrast, the success of active management involvement is dependent on other factors, such as the alignment of the CS and the parent’s businesses and strategy [37]. More research on the effects of management involvement is needed to understand its impact.

The *entity* dimension describes how the CS is structurally separated. Many studies do not define the separation in detail. We found that it can range from full integration and acting inside the incumbents’ traditional structures [29] to fully extracting it into its separate legal entity with only a few structural linkages [10]. But the entity dimension is not mappable on a one-dimensional scale. There is the idea of a safe space where the CS can act relatively freely, although not structurally separated [17]. Some structures link the CSs and the incumbent via an intermediary unit, such as an institutionalized incubator or a tech hub [19, 29]. These units themselves can be separated or integrated. The dimension entity also evolves as the CS matures. Some CS begins at a provided safe space and gets separated as it grows [26].

Branding describes an apparent external linkage to the incumbent. The association with the incumbent can evoke trust and increase credibility [17, 42]. Joint branding also simplifies joint marketing [19]. Associated branding might also increase the incumbents’

perceived dynamism and creativity [22]. However, branding wasn't a focus in these studies, and future research should consider brand research to assess its effects.

Although it is strongly linked with the entity dimension, the research we found investigated *location and facilities* separately. Some incumbents construct specialized buildings to facilitate their CS programs [29]. Partially changing locations is also used to create safe spaces and underline a new working mode for time-bound programs [8]. There could be downsides to separating the CS from the location of the incumbent, as they might loosen their relationship [17].

Program management considers that some incumbents embed CS undertakings in structured programs [19, 22]. How this management effects CSA is not described by the identified CS governance literature.

5.2 Processes and Operations

The *resources* dimension includes the resources offered and shared by the parent. This includes financials and materials, although the papers did not specify financing models extensively. Some studies describe that capital can be project-based, budget-based, granted loans, or originate from external funding sources [8, 13, 22, 29, 42]. This dimension is not limited to financial resources; it includes intangible resources like data [8] and tangible resources like equipment and infrastructure [22]. Besides the following mechanism, this dimension also encompasses resources the incumbent uses, such as their machines [22]. Furthermore, this includes human resources in the form of a workforce. In this case, the CS is either (partially) staffed by personnel from the incumbent or the CS can cooperate with the incumbents' staff [8, 13, 17, 37]. Other aspects mentioned are marketing resources like access to markets or the incumbents' network [19].

The *services* dimension encompasses a more formalized provision of resources and services. Just as the resources dimension, it includes tangible resources. In this case, these are assets provided as a service as part of a CS unit or a program [13, 17, 42]. The dimension also includes field services [42], legal services [42], human capital [8, 19, 29], and specialized facilities such as office space [22, 26, 29]. A considerable part of the services dimension involves mentoring and coaching [8, 13, 19, 22, 29].

The *structured program* dimension addresses whether firms embed the innovation process's ideation, development, and execution into a formal process. It also involves the development of ideas and whether they emerge naturally or from a structured approach. Incumbents use institutionalized accelerator programs or other innovation programs to formally assist in developing innovation [10, 19]. Nevertheless, how these programs actually interfere with CSA remains unclear.

Decision processes describe how, where and who makes decisions, involving both formal decision processes and the CSs' ability to decide independently. The authors find that rigid bureaucracy affects CSs performance negatively [13].

Metrics and KPIs describe how incumbents track CS progress. As Richter et al. [22] put it: "A company investing in such a program will likely require some evidence of return on investment which goes beyond existing accelerator metrics..." They also mention "Innovation KPIs" but do not describe the details of their function. This dimension also addresses incentive schemes for CS managers. Yang [41] finds that an incentive scheme that balances financial and strategic goals has a positive influence on a CSs performance.

Although these mechanisms effect the planning autonomy, it is unclear to what extent and in which configuration.

Scouting and selection define the process of finding and choosing innovations to pursue. This dimension also includes established scouting and selecting outside-in startups [17]. *Events* can be a part of the previously defined selection process. E.g., in the form of a demo day. They also support team-building, combine different CS initiatives, and help sophisticate a network [13, 26].

Confidentiality addresses how the CS and the incumbent share information. Although identified by Richter et al. [22] as a common feature, it is not clear how CSA is affected.

5.3 Relational Mechanisms

The dimension of *collaboration and communication* describes qualitative aspects of the collaboration between the CS and the incumbent [13, 29]. The studies identified direct access to decision-makers as a critical success factor, which goes hand in hand with the findings for the management dimension. But also, collaboration with the incumbents' employees as partners or experts is essential [13, 17]. The participants of the study by Gutmann et al. [13] recognized that ongoing cooperation was hard to establish as the incumbents' employees were not committed enough in the long term. This shows a negative effect of low CSA.

Furthermore, the articles identified the *interplay and networking* between innovation initiatives as essential. The incumbent can establish relationships between several CSs by offering a collaboration platform [13, 17]. This network facilitates an interplay between programs to enable overarching strategic innovation goals [19].

Values and culture describe how the corporate culture influences the work at the CS and could mean a culture transfer, e.g., by employing incumbent personnel at the CS. The studies generally perceive this circumstance as harmful to the CS's success [19, 26]. The studies suggest that an entrepreneurial culture that enables creativity, openness, and individual responsibility is beneficial [19, 22].

Last but not least, Selig et al. [26] outline how creating entrepreneurial *role models* that have experience and can communicate best practices, positively affects CS employees.

Table 3. Corporate startup governance mechanisms on autonomy

Mechanisms	No	Sources	Autonomy relation
Structures			
Management	7	[8, 17, 22, 26, 29, 37, 41]	Unclear
Entity	6	[10, 17, 19, 22, 29, 37]	Structural
Branding	4	[17, 19, 22, 42]	Unclear
Location	3	[8, 17, 29]	Structural

(continued)

Table 3. (continued)

Mechanisms	No	Sources	Autonomy relation
Program Mgt	2	[13, 22]	Structural
Processes and Operations			
Resources	10	[8, 13, 17, 19, 22, 23, 26, 29, 37, 41, 42]	Planning; Operational
Services	9	[8, 13, 17, 19, 22, 23, 26, 29, 41, 42]	Operational
Structured Programs	8	[10, 13, 17, 19, 22, 23, 26, 29, 37]	Unclear
Decision Processes	6	[10, 13, 17, 19, 26, 41]	Planning; Operational
Metrics	2	[22, 41]	Planning;
Scouting	2	[17, 22]	Unclear
Events	2	[13, 26]	Unclear
Confidentiality	1	[22]	Operational
Relational Mechanisms			
Collab. And Comm	5	[13, 17, 19, 22, 29]	Operational
Interplay and Netw	4	[13, 19, 26, 29]	Unclear
Values and Culture	4	[17, 19, 22, 26]	Planning; Operational
Role Models	1	[26]	Unclear
Autonomy			
Structural Aut	7	[10, 17, 19, 22, 29, 37, 41]	
Planning Aut	6	[10, 19, 22, 26, 37, 41]	
Operational Aut	5	[10, 17, 22, 37, 41]	

5.4 Autonomy

The papers cover *structural autonomy* mainly through structural mechanisms such as *entity*, *equity*, or *location and facilities* described above. They also define structural autonomy as being “structurally separated” [37]. As described in the theory section, this direct link was expected due to its nature. However, this is certainly not the case when it comes to *management dimension*. As Waldkirch et al. [37] analyze extensively, management involvement influences *structural* and *planning autonomy*. Management mechanisms seem to play a unique role in granting autonomy to CS, but the research is still fuzzy. Except for the ability to free decision-making and management interventions, we could not find any direct link between the identified governance mechanisms and CS *planning autonomy* [19, 22, 37]. Yang [41] collects data about the CS’ *planning autonomy* without asking about specific governance mechanisms. The articles primarily collect data on *planning autonomy* by asking questions about setting the CS’ own goals or being able to develop their strategy independently [10, 41]. How the programs obtain these abilities from a governance perspective is uncertain.

While Waldkirch et. al. [37] define *operational autonomy* as “...the extent to which the venture’s management team is responsible for the venture’s operations”, Garrett and

Covin [10] describe *operational autonomy* as "...the extent to which a venture has structural or process linkages back to its parent firm". From a governance perspective, these are interpreted as *structural mechanisms* instead. Yang [41] describes *operational autonomy* as hiring anyone the CS needs or making investment decisions independently. The separation of *structural, planning, and operational autonomy* remains unclear. Exact governance mechanisms that influence *operational autonomy* are missing from the analyzed literature.

6 Discussion and Future Research

Although most studies focus on operational governance, the research on governance mechanisms for CS is vast. Nevertheless, how incumbents manage CSA from a governance perspective seems to be inconsistent. The autonomy dimensions found in the literature are defined inconsistently by researchers. Likewise, how governance mechanisms institutionalize these autonomy aspects varies just as much, as there is no clear link between the applied governance dimensions and the investigated autonomy dimensions such as planning and operational autonomy. Even though we can map some of the governance mechanisms to a respective autonomy dimension with the current state of research, as shown in Table 3, there is no definitive way to build a mechanism framework for governing CSA. Furthermore, there is an imbalance of research focusing on operational governance and a strong focus on structural autonomy. To sum up, CS practitioners would benefit from a clear conceptualisation of governance models for CS and an evaluation of the associated performance effects. The following sections discuss the findings for CS governance and its relation to CSA (RQ1, RQ2) while integrating possible pathways for future research (RQ3).

6.1 Corporate Startup Governance Model

We describe CS governance mechanisms systematically and identify gaps by mapping the existing CS governance mechanisms to an established governance framework (RQ1) [36]. Governance mechanisms are valuable tools for incumbents in designing CS, and the mechanisms addressed in research represent established governance dimensions.

The current body of knowledge comprehensively investigates *structures, processes and operations*. However, some research is still needed to operationalize these mechanisms into a comprehensive model for quantitative studies. Additionally, how incumbents can manifest different characteristics of these mechanisms is still vague. To exemplify this knowledge and further substantiate the model, research should ask the following questions: (1) Which precise characteristics are specific governance mechanisms adopting in a CS context? (2) How do these forms influence the success of the CS?

Although the questions above are just as relevant for the *relational dimension*, more research is needed to define its mechanisms conceptually. There is still little research on its mechanisms from a governance perspective, although governance research might find these answers in different research streams. Therefore, we propose an additional research question for this dimension: (3) Which relational mechanisms can be extracted from the expanded research on the relationship between incumbents and CSs?

As the current research stream of CS governance still seems to be a niche, future research should consider that the developed model might not be comprehensive. Therefore, studies should further explore additional mechanisms and their forms; hence the fourth research question addressing CS governance is: (4) Which other CS governance mechanisms do incumbents utilize?

The described research agenda guides future research in building CS governance frameworks, enabling incumbents to establish CSs systematically and fostering explorative innovation.

6.2 Governing Autonomy

Research indicates that balancing autonomy substantially affects CS success [5, 18, 37]. The review presented, shows that how incumbents manage CSA from a governance perspective is discussed controversy (RQ2). The papers conceptually separate autonomy in its *structural*, *planning*, and *operational dimensions*, but the dimensions are conceptually defined inconsistently. Thus, we need to understand this concept in more detail to enable incumbents to steer autonomy actively. Therefore, we propose the following research question for future research: (5) How are *structural*, *operational*, and *planning autonomy* conceptually differentiated and defined from a governance perspective?

While in the case of structural governance and autonomy, the relationship between the dimensions is relatively well understood, this is not the case for the other two dimensions. Future studies need to answer the following research questions to close this gap: (6) How do CS governance and CSA relate? (7) How can incumbents manage CSA from a governance perspective?

7 Conclusion

This systematic literature review has built a preliminary governance framework for CSA. This might help practitioners in the context of CS to analyze the governance models available so far. To assure generalizability and applicability, we incorporated mechanisms found for inside-out and outside-in types of CSs, which we oriented on the well-established typology by Weiblen and Chesbrough [39]. Additionally, we mapped the mechanisms to the established governance dimensions: *structure*, *processes and operations*, and *relational mechanisms*. Designing governance mechanisms for CSs is always a challenge when it comes to balancing autonomy, and therefore we extracted and mapped how these mechanisms represent or influence the respective autonomy dimensions if applicable. In doing this, we systematically identify relevant research gaps that are missing to sophisticate the CS governance framework. Furthermore, we laid out a research agenda on the interplay of CS governance and CSA, as these constructs are intertwined, as shown by this review.

This research provides implications for academia and practice. Our model provides a basis to build on for future research. As most CS research is still exploratory, researchers need models suitable for quantitative research methods, and our model provides a possible foundation for this. Furthermore, our model provides a framework of governance mechanisms for CS and their relation to CSA. These findings fill the gap that prior

research has identified, as the evidence of current studies on CSA is contradictory [19, 37]. Finally, we provide a roadmap for further studies which enables researchers to investigate governance mechanisms and their impact on autonomy in more detail.

We can also derive relevant findings for practice. As described in the introduction incumbents still struggle designing their CS initiatives, and our research provides an overview of the possible mechanisms that studies have found to be effective. We offer a framework incumbents can apply to assess their CS design. Naturally, more research is needed, and incumbents must consider other aspects like their strategies to design their CSs confidently. Our model provides a first orientation in this regard. Finally, the model can be applied by corporates that are just starting out their CS initiatives and helps guiding the building process by providing a clear structure of mechanisms that are implemented in practice.

References

1. Blank S. What's A Startup? First Principles (2010). <https://steveblank.com/2010/01/25/whats-a-startup-first-principles/>. Accessed 18 May 2022
2. Boston Consulting Group, Overcoming the Innovation Readiness Gap (2021)
3. Burgers, J.H., Jansen, J.J.P., Van den Bosch, F.A.J., Volberda, H.W.: Structural differentiation and corporate venturing: the moderating role of formal and informal integration mechanisms. *J. Bus. Ventur.* **24**, 206–220 (2009). <https://doi.org/10.1016/J.JBUSVENT.2009.01.006>
4. Corvello, V., Steiber, A., Alänge, S.: Antecedents, processes and outcomes of collaboration between corporates and start-ups. *Rev. Manag. Sci.* (2021). <https://doi.org/10.1007/s11846-021-00510-8>
5. Covin, J., Garrett, R., Kuratko, D., Shepherd, D.: Short leash or long leash? Parenting style, initial strategic clarity, and the development of venture learning proficiency. *J. Bus. Ventur.* **35** (2020). <https://doi.org/10.1016/j.jbusvent.2019.105951>
6. Crișan, E.L., Salanță, I.I., Beleiu, I.N., Bordean, O.N., Bunduchi, R.: A systematic literature review on accelerators. *J. Technol. Transf.* **46**, 62–89 (2019). <https://doi.org/10.1007/s10961-019-09754-9>
7. Dushnitsky, G., Shapira, Z.: Entrepreneurial finance meets organizational reality: comparing investment practices and performance of corporate and independent venture capitalists. *Strateg. Manag. J.* **31**, 990–1017 (2010)
8. Gamber, M., Kruff, T., Kock, A.: Balanced give and take—an empirical study on the survival of corporate incubators. *Int. J. Innov. Manag.* **24** (2020). <https://doi.org/10.1142/S1363919620400058>
9. Garidis, K., Rossmann, A.: A framework for cooperation behavior of start-ups. *J. Small Bus. Enterp. Dev.* **26**, 877–890 (2019). <https://doi.org/10.1108/JSBED-04-2019-0125>
10. Garrett, R.P., Covin, J.G.: Internal corporate venture operations independence and performance: a knowledge-based perspective. *Entrep. Theory Pract.* **39**, 763–790 (2015). <https://doi.org/10.1111/etap.12059>
11. Garrett, R.P., Neubaum, D.O.: Top management support and initial strategic assets: a dependency model for internal corporate venture performance. *J. Prod. Innov. Manag.* **30**, 896–915 (2013). <https://doi.org/10.1111/JPIM.12036>
12. Guardiet, T., Oreschenko, A., Wawers, H.: Success factors of corporate accelerators. *Bus Entrep J* **11**, 1–25 (2022). <https://doi.org/10.47260/bej/1111>
13. Gutmann, T., Maas, C., Kanbach, D., Stubner, S.: Startups in a corporate accelerator: what is satisfying, what is relevant and what can corporates improve? *Int. J. Entrep. Innov. Manag.* **24**, 443–464 (2020). <https://doi.org/10.1504/IJEIM.2020.110098>

14. De Haes, S., Van Grembergen, W.: IT Governance and its mechanisms. *Inf. Syst. Control J.* **1**, 27–33 (2004)
15. Hausberg, J.P., Korreck, S.: Business incubators and accelerators: a co-citation analysis-based, systematic literature review. *J. Technol. Transf.* **45**, 151–176 (2020). <https://doi.org/10.1007/s10961-018-9651-y>
16. Johnson, K.L.: The role of structural and planning autonomy in the performance of internal corporate ventures. *J. Small Bus. Manag.* **50**, 469–497 (2012). <https://doi.org/10.1111/j.1540-627X.2012.00363.x>
17. Kohler, T.: Corporate accelerators: building bridges between corporations and startups. *Bus. Horiz.* **59**, 347–357 (2016). <https://doi.org/10.1016/j.bushor.2016.01.008>
18. Kötting, M.: Corporate incubators as knowledge brokers between business units and ventures. *Eur. J. Innov. Manag.* **23**, 474–499 (2019). <https://doi.org/10.1108/EJIM-12-2017-0201>
19. Kötting, M., Kuckertz, A.: Three configurations of corporate innovation programs and their interplay. *Eur. J. Innov. Manag.* **23**, 90–113 (2020). <https://doi.org/10.1108/EJIM-07-2018-0142>
20. Lee, S.U., Park, G., Kang, J.: The double-edged effects of the corporate venture capital unit's structural autonomy on corporate investors' explorative and exploitative innovation. *J. Bus. Res.* **88**, 141–149 (2018). <https://doi.org/10.1016/j.jbusres.2018.01.049>
21. O'Connor, G.C., DeMartino, R.: Organizing for radical innovation: an exploratory study of the structural aspects of RI management systems in large established firms. *J. Prod. Innov. Manag.* **23**, 475–497 (2006). <https://doi.org/10.1111/J.1540-5885.2006.00219.X>
22. Richter, N., Jackson, P., Schildhauer, T.: Outsourcing creativity: an abductive study of open innovation using corporate accelerators. *Creat. Innov. Manag.* **27**, 69–78 (2018). <https://doi.org/10.1111/caim.12252>
23. Schausberger, M.: Accelerators - categorizing a new form of cooperation. In: 11th International Scientific Conference on New Challenges of Economic and Business Development. Riga, pp. 739–750 (2019)
24. Scholl, L., Hirte, R.: Incubation in multinational corporations. In: ICE/ITMC 2018 - Proceedings. IEEE, Tech Univ Munich, TUM Sch Management, Munich, Germany, pp 1–8 (2018)
25. Schuh, G., Lau, F., Dyba, C., Vogt, F.: Deriving requirements for the organizational structure of corporate incubators. In: PICMET 2018, IEEE, pp. 1–8 (2018)
26. Selig, C.J., Gasser, T., Baltes, G.H.: How corporate accelerators foster organizational transformation: an internal perspective. In: ICE/ITMC 2018 – Proceedings, New York, USA, pp. 1–9. IEEE (2018)
27. Spender, J.-C., Corvello, V., Grimaldi, M., Rippa, P.: Startups and open innovation: a review of the literature. *Eur. J. Innov. Manag.* **20**, 4–30 (2017). <https://doi.org/10.1108/EJIM-12-2015-0131>
28. Steiber, A.: Technology management: corporate-startup co-location and how to measure the effects. *J. Technol. Manag. Innov.* **15**, 11–21 (2020). <https://doi.org/10.4067/S0718-27242020000200011>
29. Steiber, A., Alänge, S.: Corporate-startup collaboration: effects on large firms' business transformation. *Eur. J. Innov. Manag.* **24**, 235–257 (2020). <https://doi.org/10.1108/EJIM-10-2019-0312>
30. Steiber, A., Alänge, S., Corvello, V.: Evaluating corporate-startup co-creation: a critical review of the literature. *Int. J. Innov. Manag.* **25** (2021). <https://doi.org/10.1142/S1363919621500730>
31. Taran, Y., Boer, H., Lindgren, P.: A business model innovation typology. *Decis. Sci.* **46**, 301–331 (2015). <https://doi.org/10.1111/DECI.12128>
32. Teece, D.J.: Business models, business strategy and innovation. *Long Range Plann.* **43**, 172–194 (2010). <https://doi.org/10.1016/j.lrp.2009.07.003>

33. Teppo, T., Wustenhagen, R.: Why corporate venture capital funds fail - evidence from the European energy industry. *World Rev. Entrep. Manag. Sustain. Dev.* **5**, 353 (2009). <https://doi.org/10.1504/WREMSD.2009.031625>
34. Tushman, M.L., O'Reilly, C.A.: Ambidextrous organizations: managing evolutionary and revolutionary change. *Calif Manage. Rev.* **38**(4), 8–30 (1996)
35. Vandeweghe, L., Sharapov, D., Clarysse, B.: The accelerator as an organizational form: a review and reconceptualization. *Acad. Manag. Proc.* **2019**, 12167 (2019). <https://doi.org/10.5465/AMBPP.2019.12167abstract>
36. Vejseli, S., Rossmann, A., Garidis, K.: The concept of agility in IT governance and its impact on firm performance. In: *ECIS 2022 Research Papers*, p. 98 (2022)
37. Waldkirch, M., Kammerlander, N., Wiedeler, C.: Configurations for corporate venture innovation: Investigating the role of the dominant coalition. *J. Bus. Ventur.* **36**, 106137 (2021). <https://doi.org/10.1016/j.jbusvent.2021.106137>
38. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: writing a literature review. *MIS Q.* **26**, 13–23 (2002)
39. Weiblen, T., Chesbrough, H.W.: Engaging with startups to enhance corporate innovation. *Calif. Manage. Rev.* **57**, 66–90 (2015). <https://doi.org/10.1525/cmr.2015.57.2.66>
40. Wulandari, A.D., Subriadi, A.P.: Preparing the Startup ecosystem for building a new startup business during pandemic: a systematic literature review. *Int. J. Electron. Commer. Stud.* **13**, 137–160 (2022). <https://doi.org/10.7903/IJECS.1996>
41. Yang, Y.: Bilateral inter-organizational learning in corporate venture capital activity: governance characteristics, knowledge transfer, and performance. *Manag. Res. Rev.* **35**, 352–378 (2012). <https://doi.org/10.1108/01409171211222278>
42. Zheng, X., Feng, Y., Feng, Y., Yu, Y., Meng, Y., Pan, Y.: Research on influencing factors of incubating firms' performance by incubators of large companies. In: *ICSSSM 2019*. IEEE, Shenzhen University, College of Management, Shenzhen, Peoples R China, pp. 1–6 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Exploring the Finnish Impact Investing Ecosystem: Perspectives on Challenges from Technology Startups

Timo Okker¹ , Rahul Mohanani¹ , Tommi Auvinen¹ ,
and Pekka Abrahamsson² 

¹ University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä, Finland
falsetimookker@gmail.com, {rahul.p.mohanani,
tommi.p.auvinen}@jyu.fi

² Tampere University, Kalevantie 3, 33100 Tampere, Finland
pekka.abrahamsson@tuni.fi

Abstract. The increasing significance of social and environmental impact within the technology startup business sector has garnered attention. Previous research has explored impact investing and related themes in the startup context. However, despite the growing interest in this area, a noticeable gap exists in research addressing impact investing ecosystems (IIE) and ecosystem-related challenges and advantages specifically within the technology field. This study endeavors to fill this gap by examining organizations within the Finnish IIE, bridging the divide between current industry practices and academic research. This study employed an interview-based approach, featuring thirteen interviewees representing eleven participating organizations. These interviews followed a semi-structured format, with all interviewees holding roles closely linked to the technology startup context within the Finnish IIE. Utilizing the thematic synthesis approach, this research aims to elucidate the perceived challenges faced by technology startups operating within the IIE. The findings of this study underscore the diversity and multiplicity of challenges confronting startups within the IIE, spanning various functions and operations, as well as the existing financial structures. Furthermore, this study puts forth recommendations for mitigating these perceived challenges and suggests potential avenues for future research within this domain.

Keywords: Impact investing · Impact investing ecosystem · Challenges · Software startup

1 Introduction

Impact investing has surged in popularity in recent years, garnering increasing attention from both practitioners and scholars as they explore opportunities to harmonize social and environmental progress with economic gains [1]. While impact investing has firmly established itself as a viable investment strategy across various industries, its integration into the realm of information technology (IT) remains notably underrepresented in

2 information systems (IS) research [2]. The nexus between IT and impact investing has received limited scholarly attention, with only a handful of studies addressing this intersection [2–4]. Consequently, there remains a paucity of comprehensive research linking IT and the impact investing paradigm, as well as investigations into the practical implementation of impact investing within IT organizations.

Given that startup companies have been important innovation drivers within IT business for a long time [5], and the evident capacity of impact investing to contribute to environmental and societal challenges, it is imperative to delve deeper into the intersection of impact investing and IT startup research. Further, ecosystem research has become an important paradigm for both, impact investing and startup research. For instance, several studies have creditably described the characteristics of regional startup ecosystems and the barriers to ecosystem growth [6–8], and part of studies concentrate on IT and software startups [7, 9]. Despite this emphasis, there is a prominent shortage of research concerning advantages and disadvantages of technology startup ecosystems driven by the impact investing paradigm.

This study contributes to increase the knowledge by building up on existing impact investing ecosystem (IIE) research and empirical findings. This study defines IIE as a system which constitutes of separate interconnected actors operating in the same immediate environment. The study illustrates perceived challenges which retard the viability and evolution of IIEs to avoid known impediments of IIEs and foster processes and instruments in IT startups.

The data acquisition method employed in this investigation involved semi-structured interviews. The study encompasses a cohort of eleven informant organizations within the Finnish IIE, involving thirteen interviewees. The primary contribution of this study lies in the identification and description of challenges specific to technology startups operating within the Finnish IIE. Interestingly, several challenges resonate also to impediments perceived in the developing countries. As such, the study seeks to bridge extant bodies of knowledge pertaining to IIE theories and established startup ecosystem theories. This newfound knowledge has multifaceted utility, serving as a resource for informing novel impact initiatives, stimulating further research in this domain, and serving as a practical tool for averting common pitfalls in startup management. Study is multidisciplinary in nature by addressing research questions valuable for both IS and business study traditions. Moreover, given the nascent state of impact investing research within the fields of IS and IT, and the conspicuous dearth of understanding regarding its theoretical and practical applicability therein, this study contributes to narrowing this knowledge gap.

To address the overarching objectives of this paper, the following two research questions (RQ) were appointed: **RQ1:** *What are the most salient IIE-related challenges confronting technology startup enterprises?*; and **RQ2:** *How can these IIE challenges, specific to technology startups, be effectively mitigated?*

The paper is organized as follows: In Sect. 2, we explore the existing research related to IIE. Section 3 consolidates insights from previous studies, encompassing both challenges observed within IIE and those identified in the context of startup ecosystems. Section 3 provides an in-depth exploration of our chosen research methodology. Moving on to Sect. 4, we present the outcomes and findings of our study. In Sect. 5, we

engage in a comprehensive discussion of the implications stemming from these results. Finally, Sect. 6 serves as the culmination of our paper, where we present our primary conclusions.

2 Background

2.1 Impact Investing Ecosystem

IIE research has its roots in traditional business ecosystem research and has witnessed significant growth in recent years. Previous studies have explored IIE from various perspectives, including a general overview [10, 11], market-centric viewpoints [12], and regional analyses [13–15]. Within the broader context of impact investing research, IIE has emerged as a prominent research stream, with prior studies identifying three primary areas of focus: market growth issues, capital supply concerns, and investment readiness matters. Established theoretical frameworks and methodologies, such as network or actor-network-based theories [16, 17] and the theory of change [18], have been proposed to elucidate the impact investing paradigm. Numerous studies underscore the importance of identifying and examining the processes of key organizations and major stakeholders [10, 16]. Based on the existing body of research, the roles and functions within the impact investing network emerge as a noteworthy research theme within IIE.

The entrepreneurial ecosystem approach has been introduced to investigate IIE as self-sustaining systems comprising distinct interacting components. This perspective underscores the significance of assessing the current ecosystem to enhance comprehension of critical attributes, including enabling actors, challenges, and opportunities. Additionally, it integrates the conventional entrepreneurial ecosystem approach with the established OECD Social Impact Investment Framework to formulate the IIE Framework. This proposed framework encompasses six core domains: policy, markets, human capital, culture, support, and finance. Furthermore, several supplementary aspects complement the primary domains within this novel framework [19].

Additionally, IIE research has underscored the significance of locality, given notable regional disparities among impact investing communities [11, 15]. These distinctions necessitate thorough consideration in IIE research. While impact investing has historically gained traction and proven most successful in European and North American markets [18], evident barriers impede its growth in specific geographical regions [11, 14]. These regional variations call for more nuanced investigations, tailored to diverse cultural and legislative contexts. Consequently, further research into regional differences within impact ecosystems is imperative. Although scholars have increasingly emphasized studies within their respective regions [13, 14, 19], there remains a need for additional research on regional aspects. Furthermore, cross-country research endeavors have aimed to uncover and comprehend regional nuances and disparities in IIE across diverse economic and cultural domains [13, 15].

2.2 Challenges in IIE

Previous research has identified five primary categories of challenges within IIEs: legal and regulatory compliance, positioning within modern investment portfolios, underdeveloped infrastructure, limited investment opportunities, and a shortage of human capital for impact strategy management [20].

A significant concern revolves around the ambiguity surrounding the term “impact investing”. It lacks a universally accepted definition and is used inconsistently [21, 22], further compounded by divergent terminology employed by various IIE stakeholders due to their distinct professional backgrounds [23]. This discrepancy leads to communication issues where different practitioners may refer to different concepts when discussing impact investing.

Moreover, existing findings also highlight the formidable challenges associated with impact measurement and underscore issues related to transparency and credibility within impact funds [21]. Additionally, previous research underscores the burden on organizations to demonstrate social impact, coupled with a deficiency of tools for reporting impact outcomes [23]. Existing literature has identified numerous challenges and barriers that hinder the efficiency and impede the progress of IIEs. Disparities in the distribution of impact investing markets have resulted in certain regions being overshadowed within the global landscape. The absence of market enablers, notably government support, contributes to hindered and unequal opportunities in specific areas [11, 14]. Furthermore, the dearth of intermediary structures, coupled with high transaction costs and a deficiency in essential business skills [23], collectively serve as impediments for social enterprises.

The Ukrainian business community views impact investing primarily as a political and social endeavor, downplaying its commercial significance [14]. Interestingly, it has been observed that barriers, such as inadequate government support, impact the development of IIEs not only in developing nations with immature financial infrastructures but also in industrialized countries like Germany. The literature suggests that uncertain income models pose challenges to social enterprises due to discrepancies between their operations and inflexible public welfare funding, conflicts among various funding sources, and persistent market failures [23]. While traditional business ecosystems are typically perceived as self-sustaining systems [24], research findings underscore the essential role of public sector interventions in fostering the development and expansion of impact investing and IIEs [25, 26]. Consequently, the overall immaturity of the financial landscape and a lack of adequate public administration can be considered significant weaknesses for IIEs.

It's crucial to recognize that impact investing and its associated processes are in a constant state of evolution. Consequently, some of the challenges identified in prior research may have diminished in significance in the present landscape.

2.3 Startup Ecosystem Challenges

Existing research has identified a range of overarching challenges associated with startup businesses, encompassing financial constraints [27], shortages in human resources, deficient support mechanisms, and an inadequacy of conducive environmental factors [28]. Furthermore, another study specifically examined key challenges encountered during the

early stages of startups, concluding that these challenges predominantly pertain to market dynamics, financial viability, team dynamics, and product development. It also emphasizes that in addition to the frequently cited risks related to market and finances, there are noteworthy concerns surrounding the motivation of project teams and the constraints imposed by limited time [29].

While the existing research primarily relies on case studies conducted within domestic startup ecosystems with distinct markets, the core challenges remain consistent. For instance, in the Hungarian startup ecosystem, significant challenges revolve around securing financing, penetrating the market, and addressing distribution channel limitations [8]. Similarly, an investigation into Iran's startup landscape highlights challenges related to financing, human resource management, and uncertainties encompassing the market, platform, and team dynamics [7]. In the Israeli software startup ecosystem, notable challenges include cultural disparities, time zone differences, language barriers, a technology-centric approach at the expense of marketing, a dearth of domestic markets, and an inexperienced workforce [6]. A study focused on the Indian startup ecosystem underscores impediments related to market entry, hiring qualified personnel, navigating a complex and bureaucratic regulatory environment, in addition to some region-specific challenges [30]. Albeit comparing the ecosystems from different regions is challenging, existing research reasonably accents important challenges which are characteristic for all startup ecosystems such as finance challenges, lack of human resources and market uncertainty.

3 Methodology

In terms of the epistemological paradigm, this study aligns with interpretive qualitative research. To enhance the relevance of the findings and to gain an in-depth understanding of the chosen phenomenon, we chose an interview-based research approach to answer our RQs [31].

3.1 Identifying Participants

In selecting organizations for this study, it was essential to maintain research focus [32]. We included eleven organizations within the IIE, comprising both technology startups and key stakeholders. Selection criteria were as follows: organizations needed to have a clear connection to impact investing, either as a practitioner or stakeholder, demonstrate transparent and recognizable operations, and exhibit visible impact investing activities.

Notably, this study did not restrict organizations based on their roles within the IIE. Instead, the selection aimed to encompass various organization types and stakeholders, such as startup companies, private and public investor organizations, government governance entities, and support organizations. These organizations mainly operate in Finland but may also engage in international impact investing markets or prioritize internationalization. The selection process involved researchers' knowledge of the market and direct contact with the chosen organizations. Further details about the case organizations can be found in Table 1.

Table 1. Informant organizations.

Organization	Role	Sector
Finnfund ¹	Financier	Public
Osuuspankki	Financier, Asset management	Private
Organization 3	Accelerator, Financier	Private
Organization 4	Financier, Asset management	Private
Organization 5	Financier, Consulting	Private
Organization 6	Startup	Private
Business Jyväskylä	Incubator	Public
Organization 8	Startup, Consulting	Private
Geego Kids Oy	Startup	Private
Wointi Oy	Startup	Private
FiBAN	Consulting	Private

3.2 Data Acquisition and Analysis

The data for this study was acquired through in-depth semi-structured interviews with individuals representing eleven different organizations within the Finnish IIE. A total of thirteen interviews were conducted between 2020 and 2022. Two informants were interviewed from informant organizations 2 and 7, while the remaining cases featured one informant each. The empirical data for this study partly originated from the interview data utilized in previous research [26]. Previously unanalyzed portions of these interviews were analyzed further in this study. The original interviews were conducted in Finnish language only. If the original questionnaire is request, readers are encouraged to contact the authors of this study.

To enhance the validity of the findings, interview transcripts were created immediately after each interview. An iterative coding process was used to identify noteworthy observations. Multiple codes were initially defined based on the interview data and subsequently refined into themes. Thematic analysis [33] was employed to structure the data, utilizing a thematic synthesis approach. Several themes of interest had already been identified during the semi-structured interviews, as they were designed to address specific predefined research questions. These predefined themes encompassed basic information about the organization and interviewee, descriptions of impact investing, IIE actors, challenges related to the IIE, characteristics and processes of impact investing, impact targets and industry sectors, technology solutions, and the prospects of the field itself.

¹ www.finnfund.fi/en/

4 Findings

This section presents our results by addressing the main research questions (RQs). Sub-Sects. 4.1 to 4.7 cover RQ1, focusing on the key challenges faced by technology startups in the IIE domain. Sub-Sect. 4.8 deals with RQ2. The results obtained from the analysis were categorized into themes based on the identified codes. A summary of the codes, themes, and example quotations can be found in Table 2 here. Each subsection below discusses the main themes emerged from our research.

4.1 Business Model Challenges

The findings identified challenges in developing impactful business models that deliver value to end-customers. Startups face difficulties in implementing production chains for their services or products. Additionally, they encounter challenges in the areas of design and marketing. To address these challenges, startups often require support in terms of business model development from organizations specializing in the implementation of impact-oriented business models and possessing substantial expertise in marketing.

4.2 Impact Evaluation Challenges

Challenges in Defining the Impact. The definition of the concept of impact investing remains incomplete and lacks precision. Notably, within the product chain, certain components may align with and positively contribute to impact targets, while others may distinctly conflict with these objectives. This raises a broader discussion on the fundamental nature of impact and the necessity for a comprehensive definition that spans a company's entire production chain and operational processes. This discussion aligns with previous studies that have identified and explored the challenges associated with defining and implementing impact investing, as supported by prior research [22–24].

Challenges in Measuring Real Impact. Measuring the true impact of operations is a complex task, primarily involving the identification and selection of metrics that warrant monitoring and assessment. It is not always evident which metrics align with the desired impact outcomes, adding an additional layer of complexity to the measurement process.

Interpreting impact data presents significant challenges for companies lacking the requisite expertise for data analysis. While impact data may be accessible, it often exists in a format that is not readily amenable to constructing meaningful metrics and information. Moreover, the measured data may not be effectively leveraged to enhance operational processes, primarily due to the inherent challenges in measurement.

Challenges in Reporting the Impact. The pursuit of transparency in impact reporting is a complex endeavor, characterized by its challenges. These challenges are particularly pronounced in ambiguous environments, such as countries with underdeveloped infrastructures. Paradoxically, regions with the greatest need for investments often coincide with environments presenting higher investment risks. Challenge was identified in the interview with Finnfund, a Finnish development financier and impact investor, which widely operates also in developing countries providing finance to local initiatives. Thus, perceived challenges in IIE spans over a larger geographical area than the Finnish markets.

The findings of this study reveal a deficiency in both understanding and resources within companies when it comes to reporting impact in alignment with stakeholder expectations. These findings align with existing literature on the subject [24]. It's important to note that the inability to provide accurate and comprehensive impact reporting poses significant business risks as stakeholders and investors may be reluctant to engage with companies that encounter challenges in their reporting efforts.

Dilution of Impact Investing. The term “impact investing” has shown signs of dilution due to its widespread and inconsistent usage. Within the IIE, actors often employ the term incorrectly, either intentionally or unintentionally. Some actors may intentionally misuse the term for marketing or management purposes. This misuse of impact investing terminology, without a comprehensive understanding, has the potential to dilute the term and presents a significant risk of “greenwashing.”

4.3 Investment Challenges

Financial Infrastructure Challenges. Financial infrastructure challenges extend their impact across both domestic and international markets. Within the Finnish IIE, numerous public or partially public organizations engage in collaborations with international counterparts in foreign nations. However, disparities between regions and countries introduce significant impediments, given the substantial variations in jurisprudence, practices, and assumptions across these diverse contexts. These challenges can effectively deter investments made by Finnish investors to the markets of developing countries, as well as in companies operating within those regions.

On the domestic front, the financial infrastructure within the Finnish IIE faces a distinct challenge related to the availability of credible investment options for long-term product innovations. Consequently, a conundrum arises wherein traditional investors, primarily focused on startup companies, prioritize swifter growth and profit prospects over the extended developmental trajectories characteristic of such research-oriented projects.

Illiquidity of Investments. Impact investing instruments inherently possess complexity and illiquidity. These inherent characteristics render the determination of their value a challenging task, introducing a heightened level of risk compared to traditional investment instruments. Consequently, investors tend to shy away from impact investment products, thereby limiting the pool of available finance for such endeavors. These challenges associated with impact investing funds have been observed and documented in previous research [22].

Lack of Human Resources. Challenges arise in situations where startups face limitations in personnel availability to engage in the due diligence processes expected by public investors. Public investors typically necessitate a relatively comprehensive due diligence procedure before arriving at investment decisions. However, startup companies may find themselves lacking the necessary resources or capacity to adequately prepare for such processes or to effectively collaborate with potential investors.

Additionally, a broader issue lies in the overall scarcity of human resources within startup companies. Challenge is also appreciated by previous research [21]. Impact investors typically require extensive cooperation across various processes, including reporting. Startup companies often operate with relatively small teams whose roles may not be precisely defined, and individuals within the organization may be tasked with multiple responsibilities simultaneously. In such scenarios, establishing effective collaboration with investors proves to be a challenging endeavor.

Shortage of Finance. Several factors contribute to the constrained financial resources available to public sector organizations for investment in impact investing. First and foremost, many public sector entities, including municipalities and cities, grapple with budgetary deficits, creating substantial financing challenges. Secondly, the involvement of startup companies introduces a set of organizational risks that can dampen investor interest, particularly in the seed phase of startups.

Moreover, startup companies often represent relatively small-scale investment targets for traditional funds. Additionally, startup company shares tend to exhibit illiquidity, while the return on investment typically requires a longer timeframe compared to larger companies. These factors collectively render startup companies less appealing to traditional funds, leading to their exclusion from such investment vehicles.

Lastly, within the IIE, the absence of effective impact funds capable of providing financing to startup companies is a noteworthy concern. The interviewees highlighted the absence of impact investing funds in Finland during the interview period.

4.4 Legislation Challenges

Financial Regulation Challenges. Private investors encounter significant hurdles when attempting to enter the impact investing market. Impact investing instruments, notably funds, are categorized as complex and high-risk investment products, subjecting them to comprehensive financial regulations.

Stringent financial regulations place constraints on the potential investment volumes within the IIE. Presently, the creation of an investment product that could be accessible to private investors without professional investor status remains infeasible. Furthermore, the criteria for obtaining professional investor status are stringent and closely monitored by regulatory authorities. While this criterion serves to mitigate financial risks for individuals, it simultaneously restricts the pool of available funding. Additionally, entry into limited impact funds proves challenging due to the substantial minimum investment size requirements imposed.

Jurisprudence Challenges. Organizations hailing from diverse regions and cultural backgrounds often place distinct emphasis on varying legislative frameworks and case law, a phenomenon that does not always readily align or harmonize. These challenges, rooted in the divergence of legal and regulatory contexts, give rise to market risks that concern investors. Consequently, the presence of such risks diminishes the pool of potential impact-based funding available for projects in developing countries allocated by Finnish investors.

4.5 Market Challenges

Lack of Competence. The findings emphasize a significant knowledge gap among certain stakeholders within the IIE concerning their comprehension of profitable business processes and investment strategies, a trend that aligns with prior research [24]. These deficiencies in traditional investment practices exert an adverse influence on the quality of investment decisions and business strategies, thereby undermining opportunities for collaboration. This dearth of competence extends not only to the investment sector but also encompasses the available talent pool.

Furthermore, the findings illuminate a growing scarcity of specialized professionals and experts participating in innovative ventures within the software and technology startup sector. This insufficiency in human capital represents a substantial barrier to the expansion of startups operating within the IIE.

Non-marked Based Behavior. Non-market-based funding introduces additional barriers to entry for financiers who operate within market-oriented frameworks, especially within developing countries. Certain stakeholders within these markets do not align their operational and financial practices with prevailing market conditions. Such behavior introduces obstacles to the expansion of the impact investing market in developing countries by generating market anomalies and distorting the dynamics of local impact investing markets.

Furthermore, the presence of blended finance carries the potential to compromise the viability of traditional enterprises that might otherwise achieve higher profitability. Another challenge emerges when subsidized investments are predominantly directed towards relatively narrow sectors that are currently in vogue, thereby constraining growth opportunities in other potentially lucrative sectors.

Small Size of the Local Markets. Within the Finnish IIE, the limited scale of local markets and the complexity stemming from the multitude of actors present challenges to ecosystem collaboration. Consequently, numerous stakeholders tend to allocate their resources towards international markets instead of nurturing local initiatives and stakeholder networks. Such behavior diminishes the vitality of the local IIE.

4.6 SIB Challenges

Social Impact Bonds (SIBs) represent investments in experimental social projects that yield a return upon the achievement of predefined impact targets [34].

Exiguity of SIB Investments. One perceived challenge related to SIBs pertains to fundraising for impact-oriented companies or projects. The current Finnish IIE leans more towards mission-oriented objectives rather than adhering to conventional investment practices. While mission-oriented ventures pursue impactful goals, they often translate into low-risk and low-profit investments. Consequently, they struggle to attract investors and fail to mobilize the required level of investment, resulting in an insufficient volume of SIB projects.

Extensive Size and Complexity of SIBs. SIBs typically entail a comprehensive and protracted process. According to interview data, the planning and metrics development

phases of SIB projects can span several years. This extensive nature of SIBs poses challenges for many entities, including startups that typically operate with agile methodologies and rapid timelines. Previous research has characterized SIBs as complex [34]. The findings of this study underscore that the intricate governance structures and the costs associated with SIB projects render them infrequently used as a method for addressing social issues within public sector organizations. Consequently, this limits opportunities for startup companies to engage in collaborative endeavors.

4.7 Public Actor Challenges

Public and private actors within the IIE exhibit distinct management principles, posing challenges to effective collaboration. For example, startup companies operate with their own lexicon, practices, and operational frameworks, which differ significantly from those of governmental bodies and universities. Moreover, public sector organizations tend to avoid engagement with private sector brands, concentrating primarily on public administrative functions. This preference for pure public administration makes establishing efficient commercial partnerships challenging.

Public actors often lack expertise in marketing and branding of impact products and services, resulting in difficulties when coordinating these tasks in collaboration with startups. Public sector organizations often attempt to contribute to such tasks without the requisite proficiency, resulting in redundant efforts and hindrances to operations.

Competition for financial resources between public sector actors and private sector entities, such as registered associations, presents hurdles for private startups seeking financing. Existing entities may resist innovative solutions offered by private sector companies, thereby impeding the success of these companies.

Securing financing for private startups is further complicated by procurement processes that do not currently account for impact investing assets. Impact investing remains excluded from procurement specifications, and its distinctive characteristics are not factored into the process, resulting in the displacement of impact startups in procurement procedures.

Another challenge emerges from public investors' perception of impact companies as high-risk investment targets. This perception often leads to situations where financing for impact startups is either unavailable or comes at a higher cost compared to traditional companies.

4.8 Mitigation of Challenges

This section provides answers to RQs that pertain to practical implications derived from the results (RQ2). By presenting these implications, this study aims to contribute to the advancement of current research and furnish tools to assist practitioners within the IIE.

Create Impact Investing Funds. To enhance the funding of impact investing startups, a more targeted funding approach is imperative. Dedicated impact investing funds have the potential to effectively mobilize financing for startup initiatives characterized by relatively low risk profiles. Financial institutions and organizations should contemplate

the establishment of such funds exclusively dedicated to the funding of impact investing companies.

Furthermore, impact investing funds play a vital role in reducing the barriers that individual investors face when entering the impact investing markets. These funds facilitate the participation of individual investors, as they do not necessitate professional investor status for those investing through them.

Enhance Collaboration Between Public and Private Actors. Given that numerous challenges within the IIE are intricately linked to collaboration between public entities and private enterprises, it is crucial to augment cooperation and the involvement of public organizations. The findings underscore that the root causes of several challenges stem from inadequacies in competence, misunderstandings, and feeble cooperation among various IIE stakeholders. These challenges, as revealed by the findings, are primarily attributed to shortcomings within public organizations.

Enhancing collaboration can be achieved through a series of strategic actions, and we propose the implementation of impact investing training specifically tailored for public actors engaged with companies focused on impact creation. This targeted training can help bridge the competency gap and foster more effective engagement between public organizations and impact-driven enterprises.

Define the Impact. Insufficient or unclear definition of impact relates to several challenges perceived by practitioners within IIE, and the issue was mentioned in several interviews. Impact targets are still constantly defined in ambiguous ways, which leads to challenges such as weak collaboration, lack of finance and tenuous impact results.

Challenges can be tackled by creating more accurate impact analysis when defining impact targets either by resourcing people to investigate impact within the company, or by acquiring this service as a purchased service from consultation companies specialized in impact evaluation. Results also highlight impact certificates to standardize the market.

5 Discussion

This study draws several key conclusions from its analysis. Firstly, it highlights that existing IIEs do not adequately facilitate cooperation between startup companies and investors. Public organizations, including business unit organizations and private consultants, should play a more active role in fostering networking and collaboration between investors and companies, allocating sufficient resources to support these efforts.

Secondly, the study identifies challenges stemming from public organizations' limited understanding of impact investing principles and processes, which hinders the development of necessary infrastructure for impact investing and support for startup companies within the industry. Third, the lack of a precise and universally accepted definition of impact investing creates issues for impact evaluation. To address this, the study proposes the implementation of certifications to clarify and standardize the definition of impact investing and encourages companies to allocate resources to create accurate impact analysis, while also calling for academic research to provide a more comprehensive understanding of the topic.

Furthermore, the study emphasizes significant obstacles in financing startups within the IIE. It reveals a disconnect between investors and investment targets within the ecosystem, underscoring the importance of fostering productive dialogue to address perceived uncertainties. Additionally, the study advocates for the evaluation of financial regulations to align them with the urgent needs of impact investing and the startup sector. The establishment of dedicated impact investing funds is also recommended to secure funding for innovative initiatives. Moreover, the study highlights the crucial role of public investments in securing financing for startups within the IIE.

Again, despite SIBs popularity in certain sectors, results of the study indicate that SIB projects are not able to leverage significant movement among technology startups as SIBs do not prove to be attractive from the startups perspective due several significant impediments related to them. At its current state SIBs apparently remain a minority form of investment notably among Finnish based technology startups.

This study aligns with prior research on IIE challenges related to legal compliance, impact definition and reporting, impact funds, human resources, competence, and SIB projects. While some challenges resonate with issues observed in startup management research, there are unique challenges specific to IIEs. Furthermore, several challenges resonate also to the markets of developing countries as Finnish IIE actors have connections to these countries in form of development finance. Additionally, this study contributes novel insights regarding impediments faced by technology startups within IIEs, enriching the body of knowledge in this field. While primarily rooted in the IS tradition, this research also holds multidisciplinary significance, offering theoretical and practical insights relevant to fields such as management and economic sciences.

5.1 Future research

Given that several perceived impediments in IIE are related to evaluation of impact and financial infrastructure and remain rather vague in existing research, this study emphasizes further research considering these topics. For instance, research on impact evaluation processes and practices among startup practitioners and well as studies considering the comprehension of impact concepts within startup companies would be pivotal. Furthermore, due the perceived shortcomings and challenges of current SIBs, they are not considered to be effective instruments to leverage financing for innovative impact initiatives. Hence, more research on SIB in the context of technology startups is encouraged. In addition, further research related to IIE's in IS in general is important to understand the phenomenon more profoundly.

5.2 Limitations

It is crucial to acknowledge that challenges within the IIE are both numerous and multifaceted, and any single study may not comprehensively address all perceived challenges. Therefore, it is imperative to conduct further research that focuses on specific types of challenges within the IIE.

In addition, it is worth noting that synthesizing the results of this study with the existing literature on the topic is not a straightforward task. Studies related to the IIE

often have regional relevance, and their discussions are centered within specific contextual environments. While interviews provide valuable insights into delimited research subjects, their findings may not be directly generalizable.

6 Conclusions

In summary, this study endeavors to address the knowledge gap in IIE research and perceived challenges faced by technology and software startups and important stakeholders within these ecosystems. This study takes a multidisciplinary perspective to investigate perceived challenges and to provide practical implications to mitigate these challenges. The research employed a qualitative approach, utilizing semi-structured interviews for data collection. The study identifies multiple challenges encountered by various actors within the IIE, with many of these challenges remaining insufficiently addressed in previous research.

The findings of this study shed light on several challenges that are particularly salient for technology startups. Study identified multiple types of challenges within Finnish IIE which are as follows: business model challenges, impact evaluation challenges, investment challenges, legislation challenges, market challenges, SIB challenges and public actor challenges.

While issues related to impact evaluation, financing, and the availability of adequate human resources have already been recognized as challenges, this study contributes by highlighting additional challenges such as those related to business models, stakeholder dynamics, emerging market complexities, and issues specific to SIB projects. Furthermore, the study proposes three distinct perspectives for addressing the perceived challenges within the IIE, thereby enriching the body of knowledge in this field.

References

1. Agrawal, A., Hockerts, K.: Impact investing strategy: managing conflicts between impact investor and investee social enterprise. *Sustainability* **11**(15), 4117 (2019)
2. Okker, T.: Impact investing in information technology business. *Electron. J. Bus. Ethics Organ. Stud.* **27**(1), 16–28 (2022)
3. Jablonski, A., Jablonski, M.: Impact investing in digital business models. *Energies* **14**(18), 5785 (2021)
4. Gidron, B., Israel-Cohen, Y., Bar, K., Silberstein, D., Lustig, M., Kandel, D.: Impact tech startups: a conceptual framework, machine-learning-based methodology and future research directions. *Sustainability* **13**(18), 10048 (2021)
5. Baskerville, R., Ramesh, B., Levine, L., Pries-Heje, J., Slaughter, S.: Is internet-speed software development different? *IEEE Softw.* **20**(6), 70–77 (2003)
6. Kon, F., Cukier, D., Melo, C., Hazzan, O., Yuklea, H.: A panorama of the Israeli software startup ecosystem. <https://ssrn.com/abstract=2441157>. Accessed 27 Aug 2023
7. Salamzadeh, A., Kesim, H.K.: The enterprising communities and startup ecosystem in Iran. *J. Enterprising Commun. People Places Global Econ.* **11**(4), 456–479 (2017)
8. Jaki, E., Molnar, E., Kadar, B.: Characteristics and challenges of the Hungarian startup ecosystem. *Budapest Manage. Rev.* **50**(5), 2–12 (2019)

9. Tripathi, N., Seppänen, P., Boominathan, G., Oivo, M., Liukkunen, K.: Insights into startup ecosystems through exploration of multi-vocal literature. *Inf. Softw. Technol.* **105**, 56–77 (2019)
10. Alijani, S., Karyotis, C.: Coping with impact investing antagonistic objectives: a multistakeholder approach. *Res. Int. Bus. Financ.* **47**, 10–17 (2019)
11. Islam, S.: M: Impact investing in social sector organizations: a systematic review and research agenda. *Account. Financ.* **62**(1), 709–737 (2022)
12. Tekula, R., Andersen, K.: The role of government, nonprofit, and private facilitation of the impact investing marketplace. *Public Perform. Manag. Rev.* **42**(1), 142–161 (2019)
13. Acevedo, J., Wu, M.: A proposed framework to analyze the impact investing ecosystem in a cross-country perspective. *Rev. Eur. Stud.* **10**(4), 87–113 (2018)
14. Bukharina, L., Onyshchenko, O.: Development of the impact investing ecosystem in Ukraine. *Investment Manage. Financ. Innovations* **16**(3), 217–228 (2019)
15. Roundy, P.: Regional differences in impact investment: a theory of impact investing ecosystems. *Soc. Responsib. J.* **16**(4), 467–485 (2019)
16. Michelucci, F.: Social impact investments: does an alternative to the anglo-saxon paradigm exist? *Voluntas* **28**(6), 2683–2706 (2017)
17. Tsujimoto, M., Kajikawa, Y., Tomita, J., Matsumoto, Y.: A review of the ecosystem concept — towards coherent ecosystem design. *Technol. Forecast Soc. Change* **136**, 49–58 (2018)
18. Jackson, E.: Interrogating the theory of change: evaluating impact investing where it matters most. *J. Sustain. Finance Investment* **3**(2), 95–110 (2013)
19. Acevedo, J., Wu, M.: The impact investing ecosystem in Japan and Singapore. In: 24th International Scientific Conference on Economic and Social Development, pp. 120–148. Faculty of Management University of Warsaw, Warsaw (2017)
20. Ormiston, J., Charlton, K., Donald, M., Seymour, R.: Overcoming the challenges of impact investing: Insights from leading investors. *J. Soc. Entrepreneurship* **6**(3), 352–378 (2015)
21. Combs, K.: More than just a trend: the importance of impact investing. *Corp. Financ. Rev.* **18**(6), 12–18 (2014)
22. Höchstädter, A., Scheck, B.: What’s in a name: an analysis of impact investing understandings by academics and practitioners. *J. Bus. Ethics* **132**(2), 449–475 (2015)
23. Glänzel, G., Scheuerle, T.: Social impact investing in Germany: current impediments from investors’ and social entrepreneurs’ perspectives. *Int. J. Voluntary Nonprofit Organ.* **27**(4), 1638–1668 (2016)
24. Peltoniemi, M., Vuori, E.: Business ecosystem as the new approach to complex adaptive business environments. *Proc. eBusiness Res. Forum* **2**(22), 267–281 (2004)
25. Mendell, M., Barbosa, E.: Impact investing: a preliminary analysis of emergent primary and secondary exchange platforms. *J. Sustain. Financ. Investment* **3**(2), 111–123 (2013)
26. Okker, T., Auvinen, T., Abrahamsson, P.: Characteristics and actors of impact investing ecosystem from the perspective of technology startups. In: ICSOB-C 2022 International Conference on Software Business: Companion Proceedings (PhD Retreat, Posters and Industry Track) of the 13th International Conference on Software Business, CEUR Workshop Proceedings, Bolzano, Italy (2023)
27. Colombo, M.G., Piva, E.: Strengths and weaknesses of academic startups: a conceptual model. *IEEE Trans. Eng. Manage.* **55**(1), 37–49 (2008)
28. Salamzadeh, A., Kesim, H.K.: Startup companies: life cycle and challenges. In: 4th International Conference on Employment, Education and Entrepreneurship (EEE), Belgrade, Serbia (2015)
29. Giardino, C., Bajwa, S.S., Wang, X., Abrahamsson, P.: Key Challenges in Early-Stage Software Startups. In: Lassenius, C., Dingsøyr, T., Paasivaara, M. (eds.) XP 2015. LNBIIP, vol. 212, pp. 52–63. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18612-2_5

30. Korreck, S.: The Indian Startup Ecosystem: Drivers, Challenges, and Pillars of Support, vol. 210. ORF Occasional Paper (2019)
31. Klein, H.K., Myers, M.D.: A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Q.* **23**(1), 67–93 (1999)
32. Baxter, P., Jack, S.: Qualitative case study methodology: study design and implementation for novice researchers. *Qual. Rep.* **13**(4), 544–559 (2008)
33. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement, pp. 275–284. IEEE
34. Warner, M.: Private finance for public goods: social impact bonds. *J. Econ. Policy Reform* **16**(4), 303–319 (2013)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Practitioner Views on Analytics for Software Startups: A Preliminary Guide Based on Gray Literature

Usman Rafiq¹, Frédéric Pattyn², and Xiaofeng Wang¹

¹ Faculty of Engineering, Free University of Bozen-Bolzano, Bolzano, Italy
{urafiq,xiaofeng.wang}@unibz.it

² Department of Business Informatics and Operations Management,
Ghent University, Ghent, Belgium
Frederic.Pattyn@ugent.be

Abstract. Software startup companies operate under extreme conditions of uncertainty and with limited resources. These innovative companies face constant pressure to find a product-market fit, drive growth, and maintain competitive advantage. The nature of these companies makes them suitable candidates to practice analytics. Analytics can help software startups to use data in several ways e.g. make data-informed decisions, grow business, and provide value to users. However, startup founders tend to put off practicing analytics for a later time. In addition, the existing literature on startups does not provide paved paths to establish analytics in the context of startups. Therefore, to this end, we perform a gray literature review, to understand what startup practitioners say about analytics benefits and how can startups define analytics within their particular context. We utilized YouTube as a source of our data. After applying inclusion and exclusion criteria to 400 videos, we ended up analyzing 16 potentially relevant videos. We used thematic synthesis as well as quasi-statistics to analyze the data. Our results identify and report ten analytics benefits, and two key analytics practices to set up analytics in these competitive environments.

Keywords: Data-Analytics · Benefits · Practices · Software Business · Metrics

1 Introduction

Software startups are significantly contributing to making the world a better place. Today's most influential software businesses initiated their journey as a startup. Netflix, Airbnb, Uber, LinkedIn, Canva, and Slack are only a handful of instances. These small yet innovative companies are witnessed driving the economy of today's contemporary world [10]. Innovation, uncertainty, scarcity of resources, high reactivity, and time pressure are some notable characteristics that distinguish these companies from other software businesses [6]. The proliferation

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 315–323, 2024.

https://doi.org/10.1007/978-3-031-53227-6_22

of startups across the globe is continuously booming. Nevertheless, more than 90% of the startups completely fail and only 15% of those that sustain themselves get a successful exit [4,10]. This high failure alludes to how much money startups have wasted and may continue to waste. The significant reasons identified after studying thousands of startups are actually related to each other i.e. no market need and running out of cash [10].

On the other hand, analytics has become more and more prevalent for a wide range of companies including software companies(e.g. software analytics [12]). For these established companies, evidence indicates that analytics can play a pivotal role in maximizing the productivity of companies, reducing costs, helping to identify trends, and maintaining competitor advantage [11]. However, when it comes to startups, there is a lack of a comprehensive understanding of what constitutes analytics for startups and how startups can utilize it to drive success and growth. Therefore, this study fills the gap in the academic literature by attempting to understand how startups can benefit from analytics in terms of raising the odds of success, reducing uncertainties, coping with dynamic markets, and learning. Thus, the following Research Questions(RQs) are guiding our study: What benefits, related to analytics, do software startups ascertain?(**RQ1**) and What are the key practices to define analytics inside startups?(**RQ2**)

We performed a Gray Literature (GL) review [8] and collected videos as GL data source to address our RQs. We identified 16 relevant videos and then used thematic analysis and quasi-statistic to synthesize findings. Therefore, we identify and present ten opportunities that analytics can bring to startups along with two analytics practices. These results aim to help startup companies in defining the analytics setup.

2 Related Work

Despite the ever-growing significance of analytics, there is a lack of knowledge regarding what constitutes analytics for software startups and how can these companies utilize it.

A few recent studies [3–5] develop our earlier understanding of analytics for startups in terms of role of analytics in startup companies, analytics challenges for startups, and perception of startups regarding analytics. Much of the related work, in the field of software engineering, is focused on analytics about software and its associated artifacts [12]. Therefore, it still remains a challenge to translate many of existing research insights into actionable steps, especially within the unique environment of startups.

3 Research Method

We conducted a Gray Literature(GL) review [8] due to the lack of existing scholarly research and limited access to primary data. We also aimed to elicit knowledge from startup practitioners who are directly influencing novice entrepreneurs

[6]. The use of GL is not a new development in the field of Software Engineering (SE). Several studies in SE and startups (e.g. [1, 2, 6, 7]) have utilized GL, particularly selecting, web pages, blogs, videos, books, technical reports, or white papers, as data sources.

We utilized YouTube to collect GL data. Our eight search strings included “software startup”, and “analytics” and its associated terms. We expanded our search to include the first 50 search results for each search string. After applying inclusion/exclusion criteria to 400 videos, we identified 16 potentially relevant videos addressing the RQs. The final version of the dataset contained 415 min of videos (seven hours) [14], and 81574 words (181 pages).

We started our data analysis by extracting metadata and demographics of practitioners. Later, we performed thematic analysis [9] to synthesize the data, focusing on identifying recurring themes within the data. In conjunction with thematic analysis, we also applied quasi-statistics [13] method that advocates to identify the most frequently occurred analytics benefits and practices.

4 RQ1: Benefits of Analytics for Software Startups

B1: Data-Driven Decision Making

Facilitating startups to make data-driven decisions appeared as one of the key advantages characterized by several practitioners. Smart decisions, quick decisions, and informed decisions are the possible outcomes startups can achieve by utilizing analytics. For example, in the instance of GL8, the practitioner reported: *‘By understanding these metrics, data-driven business decisions can be made’*. Decisions cover a wide range of tasks in which startup founders must be interested. It includes decisions, for instance, identifying best-performing acquisition channels or identifying the type of interested users.

B2: Improving Efficiency and Focus

Startups can certainly improve their business efficiency and start focusing on things that really matter. A practitioner from [GL3] alluded: *‘you want to start using data to drive your focus’*. It is complemented by another practitioner in [GL9] in the following words: *“[Analytics] helps you really keep it there, like figure out where to start, where to focus...your efforts when you’re thinking about your product. and what to do next”*.

B3: Visibility and Realism

B3 promises comprehensive visibility of the startup as a business, and, more importantly, brings founders closer to reality. According to [GL1], visibility means *“what’s going on across our business in the corner of our eye...knowing that if something big happens we’re not going to miss it.”* On the other hand, startup founders are always in love with their ideas [4]. Here, *“analytics helps you [to] be real with yourself. Do customers actually want this?”*, added by practitioner from [GL13].

B4: Enhancing User Experience

Startups can achieve user experience enhancement by using analytics in several ways. For example, by getting in-depth user insights, improving user engagement, and maximizing user retention. The practitioner [GL3] encouraged this in the following words: “*[understand] who is the user and what are some characteristics of this user*”. Another practitioner from [GL9] goes deeper into this and explains the user understanding process: “*[Identify] what are the demographics, behavioral details, what are their needs, obstacles...you likely might have already some sort of profile of your users...*”.

B5: Fostering Data-Driven Culture

Analytics can foster a data-driven culture inside a startup. Eventually, data becomes the language that everyone speaks in the company. It is reported at length, for instance, in [GL12] in the following words: “*Want to have a culture at your startup that believes in data...that looks at the metrics all the time and that starts at the top, the CEO, and the VPs... the people who watch these numbers, who measure these numbers... And who talk about them in group meetings, who talk about them in their emails*”.

B6: Understanding and Insights

B6 promises comprehensive real-time insights to understand various actions and outcomes for a startup. It covers aspects like “*what’s happening right now*”, as the practitioner [GL1] reported. The practitioner continued explaining this in the following excerpt: “*something great, maybe we’re featured in a blog post that we didn’t expect to get a huge influx of traffic*”. A similar indication about real-time insights is furnished by [GL11] in the following words: “*it is important because obviously, you should know what state your business is in at all times*”.

B7: Detecting Growth Challenges

Analytics helps startups to detect all the possible user growth issues as well. Startups might get some customers early on but then the user growth, retention, or engagement decreases. According to the practitioner from [GL3], one apparent reason is the product-market fit. He mentions this in the following words: “*the products that have no product market, the engagement over time, for all cohorts, will go to zero*”.

B8: Team Alignment

Another noteworthy benefit that analytics can offer is team alignment. The insights obtained through analytics can make everyone on the same page. This is supported by a practitioner from [GL3], who expressed his opinion in the following excerpt: “*you want to motivate your team... use this data... So what*

you're gonna do is you're gonna set [shared] goals". Adding on top of that, while explaining the questions that analytics can help out with, the practitioner from [GL9] commented: "...and so this helps to create alignment on your team".

B9: Improving Product Usability

Startups, usually in the early stages, need to launch their products. They can assess with the help of analytics how usable the product is, how users are using it, do users understand the product, and which features are getting popular. The practitioner at [GL2] thinks that "*almost every product that's launched is unusable or highly unusable for the first three months*". That is the time to improve product usability through analytics.

B10: Supporting Product Development and Enhancement

This theme reports two perspectives. The first one is related to testing the product market fit, a fundamental activity for startups. The second one is accelerating product development through analytics. Both perspectives insist on a feedback mechanism to elicit user behavior. A practitioner from [GL13] reported: "*analytics is incredibly important... it helps you test product-market fit*"? Another practitioner from [GL11] agrees and states its use in "*building new features, launching new features, and so on*" (Table 1).

Table 1. Overview of the Identified Benefits of Analytics for Software Startups

Benefit ID	Benefit Title	Instances	Videos
B1	Data-Driven Decision Making	21	[GL2][GL3][GL4][GL5][GL6][GL7][GL8][GL9][GL11][GL13]
B2	Improving Efficiency and Focus	14	[GL3][GL4][GL7][GL9][GL11][GL14]
B3	Visibility and Realism	13	[GL1][GL3][GL5][GL8][GL11][GL13][GL16]
B4	Enhancing User Experience	24	[GL2][GL3][GL4][GL5][GL6][GL9][GL10][GL12][GL13][GL14][GL16]
B5	Fostering Data-Driven Culture	5	[GL2][GL3][GL5][GL9][GL12]
B6	Understanding and Insights	15	[GL1][GL2][GL3][GL4][GL5][GL9][GL11][GL14]
B7	Detecting Growth Challenges	4	[GL2][GL3][GL4][GL10]
B8	Team Alignment	4	[GL3][GL9][GL11][GL16]
B9	Improving Product Usability	5	[GL2][GL3][GL9][GL14][GL16]
B10	Supporting Product Development and Enhancement	9	[GL2][GL3][GL5][GL9][GL10][GL11][GL13][GL15][GL16]

5 RQ2: Practices to Define Analytics in Software Startups

5.1 Prioritize Key Metrics

The most prominent advice reported by practitioners is to **identify top-level KPIs** first. It is explicitly highlighted in 11 videos. While there exists a lot of definitions of KPI, the practitioner from [GL11] defines it as a “*set of quantitative metrics that indicate how healthy your business is doing*”. There are a plethora of metrics available to startups. However, like others, practitioner in [GL3], indicated to select one. He expressed it in the following words: “*there is usually almost only one metric that represents value for each company*”.

Thereafter, in eight videos, there are guidelines on **selecting and defining the KPI** from a variety of metrics. The practitioner in [GL1] guided in the following words: “*the one metric that matters is the metric that you choose to focus on, so that’s the metric that you’ve decided will have the biggest impact on your growth*”. Going into more details and while guiding how startups can selecting top-level KPIs, a practitioner from [GL16] commented: “*what is a number that you’re willing to bet the company on? If that number goes south. You deserve to die. And if that number goes up. You will like...you will have made a huge difference in the universe*”. Our data analysis also reveals that the business domain of a startup is an important factor in deciding the top-level KPI. It will vary from domain to domain and thus there is no silver bullet.

Later, **adding supporting metrics to top-level KPI** is considered an essential step. It is found in four videos. Some practitioners like [GL1] referred to it as “nuance” metrics while others, such as, [GL9] referred to it as secondary metrics. However, the purpose remains the same. As an example, if the selected KPI for an e-commerce startup is the number of sales then average sales or a unique number of customers will help to present the full picture with top-level KPI [GL1].

Lastly, we come across the indication of **regular monitoring of selected KPI**. Practitioners consider monitoring and taking action based on monitoring as essential as the selection itself. Commenting on this, the practitioner in [GL1] mentioned: “*if we pick KPIs and then ignore them... we’re also in trouble...if we pick and monitor our KPIs diligently but we don’t assess... everything we do and everything a whole team does around o...at the end of the day, we’re still screwed*”.

5.2 Keep Analytics Simple

This theme classifies and presents high-level codes that **strive to educate startups** on the basics of setting analytics in their companies. The first lesson practitioners communicate here is to learn that ‘less is more’. Our data analysis, based on instances found in seven videos, highlighted that some founders become overwhelmed with analytics as they attempt to model every aspect of their startup. It is apparent from the following excerpt of a practitioner [GL8] “*The point is not*

to track everything because eventually if you do try to track everything, you're just going to be... ended up in a [situation] where you're just tracking things without actually making it... decisions without actions". Another practitioner[GL16] expressed: *"Don't boil the ocean..."*. *"less is more"*, he added further.

Next, we have a very similar but critical issue, labeled as **"analysis paralysis"**. This situation occurs when a startup starts over-complicating analytics stuff e.g. selecting the best analytics tool, building a tool from scratch, thinking too much about selecting the right metrics, and putting a lot of time into looking at the data. The issue is referred to as analysis paralysis. One of the practitioners[GL1] warns startups by pointing out how to know if they are doing analysis paralysis. The practitioner reported: *"when are you spending too much time looking at the numbers? versus actually action stuff"*.

Along the same lines, **accurate estimates are not required** when a startup is using analytics. It was highlighted in four videos in different instances. For instance, the practitioner[GL5] advised it in the following excerpt: *"you're a startup. You're not going to have a lot of data to be able to do like fine-grained analysis... You may have some data, you may have other people's data, you can still draw a box. around. products"*.

The last category in this theme refers to the **adoption and focus** regarding analytics. This was presented in five videos. It states that with the passage of time, focus on KPIs and metrics change, tools change and business segments change as well when startups pivot. As an example, the practitioner[GL10] clearly emphasized: *"companies mature and grow, they start to shift their attention from the metrics that they used in the beginning stages of their business to metrics that are important later on in their business"*.

6 Conclusions and Future Work

Our research presented ten analytics benefits and two practices for software startups, drawing on experiences of startup practitioners. Primarily, our findings are particularly relevant for early-stage startups, as these companies are often hesitant to practice analytics. On the other hand, we conclude that while there is no silver bullet solution to define the top-level KPI, answering a few questions and the business domain of a startup might contribute to define it. Likewise, our results also highlight areas directly influenced by analytics. For example, the immediate impact of using analytics produces product design decisions, product engagement strategies, and enhancement of user experiences. At same time, analytics is found offering a supporting role to solve fundamental pain points of startups. It includes identifying the target customers, target market, or testing product market fit.

In the current study, we fell short of utilizing snowballing techniques to figure out more related videos such as YouTube recommendations and indications of other sources in our data. Therefore, this remains an important addition for future work. Moreover, additional work is needed to include blog posts and website data to draw a full picture of analytics inside startup companies. Therefore, we intend to take these variables into account for our immediate future work.

References

1. Chattopadhyay, S., Zimmermann, T., Ford, D.: Reel life vs. real life: how software developers share their daily life through vlogs. In: Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 404–415 (2021)
2. Christmann, M., Mayer, P., Volkamer, M.: Vision: what Johnny learns about password security from videos posted on YouTube. In: Proceedings Of The 2021 European Symposium On Usable Security, pp. 124–128 (2021)
3. Rafiq, U.: Towards understanding analytics in software startups. In: Proceedings of the 5th International Workshop on Software-Intensive Business: Towards Sustainable Software Business, pp. 31–38 (2023)
4. Rafiq, U., Melegati, J., Khanna, D., Guerra, E., Wang, X.: Analytics mistakes that derail software startups. In: Proceedings of the 25th International Conference on Evaluation and Assessment in Software Engineering, pp. 60–69 (2021)
5. Berg, V., Birkeland, J., Pappas, I., Jaccheri, L.: The role of data analytics in startup companies: exploring challenges and barriers. In: Challenges and Opportunities in the Digital Era, pp. 205–216 (2018)
6. Melegati, J., Guerra, E., Wang, X.: Understanding hypotheses engineering in software startups through a gray literature review. *Inf. Softw. Technol.* **133**, 106465 (2021)
7. Tripathi, N., Seppänen, P., Boominathan, G., Oivo, M., Liukkunen, K.: Insights into startup ecosystems through exploration of multi-vocal literature. *Inf. Softw. Technol.* **105**, 56–77 (2019)
8. Garousi, V., Felderer, M., Mäntylä, M.V.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Inf. Softw. Technol.* **106**, 101–121 (2019)
9. Cruzes, D., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement, pp. 275–284 (2011)
10. Genome, S.: Global Startup Ecosystem Report 2022. Technical Report, Startup Genome, San Francisco, CA (2022)
11. Bianchini, M., Michalkova, V.: Data Analytics in SMEs: Trends and Policies. OECD, Paris (2019)
12. Zhang, D., Han, S., Dang, Y., Lou, J.G., Zhang, H., Xie, T.: Software analytics in practice. *IEEE Softw.* **30**(5), 30–37 (2013)
13. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empir. Softw. Eng.* **14**, 131–164 (2009)
14. Usman Rafiq, Frederic Pattyn, and Xiaofeng Wang. The Dataset used for exploring benefits and practices of applying analytics in software startups. FigShare, Dataset (2023). <https://doi.org/10.6084/m9.figshare.24153333.v3>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Software Product Management



An Evaluation of the Product Security Maturity Model Through Case Studies at 15 Software Producing Organizations

Elena Baninemeh¹(✉), Harold Toomey², Katsiaryna Labunets¹, Gerard Wagenaar¹, and Slinger Jansen^{1,3}

¹ Utrecht University, Utrecht, The Netherlands
{e.baninemeh,k.labunets,g.wagenaar,slinger.jansen}@uu.nl

² Raytheon Technologies, Waltham, USA
Harold@Toomey.org

³ LUT University, Lappeenranta, Finland

Abstract. Cybersecurity is becoming increasingly important from a software business perspective. The software that is produced and sold generally becomes part of a complex landscape of customer applications and enlarges the risk that customer organizations take. Increasingly, software producing organizations are realizing that they are on the front lines of the cybersecurity battles. Maintaining security in a software product and software production process directly influences the livelihood of a software business. There are many models for evaluating security of software products. The product security maturity model is commonly used in the industry but has not received academic recognition. In this paper we report on the evaluation of the product security maturity model on usefulness, applicability, and effectiveness. The evaluation has been performed through 15 case studies. We find that the model, though rudimentary, serves medium to large organizations well and that the model is not so applicable within smaller organizations.

Keywords: software product security · software engineering security · product security maturity model

1 Introduction

“Cybersecurity is the collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment and organization and user’s assets.” [42]. It strives to ensure the integrity, availability and confidentiality of software applications. There are plenty of tools, such as firewalls and antivirus software to prevent cyber-attacks and detect security breaches. A cyber-attack is action where a person tries to penetrate another person’s computers or network for the purpose of causing damage or disruption [11]. Cybersecurity tries to prevent a cyber-attack from happening. We argue that

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 327–343, 2024.

https://doi.org/10.1007/978-3-031-53227-6_23

cybersecurity is one of the recently introduced cost factors in SPOs and that this field deserves more attention from the software business research community. During the development phase of a software product, one of the key priorities for software engineers is ensuring the fulfillment of quality and security requirements [10]. Software business has benefited from maturity models [17, 38]. Several maturity models [4] are being used by Software Producing Organizations (SPOs) to evaluate their software product and software production security. One of these models, called the Product Security Maturity Model (PSMM) that has not sufficiently been evaluated for its usefulness and applicability, so in this study, we improve this problem by evaluating the PSMM.

In the next Section, we introduce the PSMM. In Sect. 3 we reiterate the objective of this work and describe how we performed a model comparison and a holistic multiple case study at 15 organizations with a large number of small research teams.

1. In Sect. 4, we compared the PSMM with BSIMM and SAMM and discovered that the PSMM is unique in its agility and relative completeness for SPOs.
2. Secondly, we report on 15 case studies in Sect. 5, with the goal of identifying patterns in the data. We find that operational security is directly related to size of the company, but that technical product security is not dependent on a company's size.
3. With the participants in the case studies, we also evaluate the usefulness, applicability, and effectiveness of the PSMM and report on the findings from those evaluations in Sect. 6. We discovered that the model was proficient in suggesting new security practices to the participants in the case study. However, it does suffer from certain design flaws. Furthermore, in Sect. 6.1, we discussed various situational factors that were identified.

We conclude the work with a discussion about the role of maturity models as a scientific endeavor and their role in improving SPOs.

2 Introducing the PSMM

Evaluating the cybersecurity of any business is a difficult endeavor, comparing these evaluations is even more of a challenge, especially so if the evaluations were done according to different metrics. To solve this issue and evaluate whether partners were using proper cybersecurity protocols, an employee at semiconductor chip manufacturer Intel developed the “Product Security Maturity Model”¹.

The PSMM evaluates based on twenty criteria, which are split in two categories: Operational and Technical. Operational parameters in PSMM include measures of program support, staffing and resources, SDL implementation, protection from externally reported product vulnerabilities (PSIRT), adherence to product security policies and processes, security training, and efficiency of data tracking and security metrics. Technical parameters in PSMM include measures

¹ www.toomey.org/psmm/.

of software security requirements and verification, software architecture and design reviews, threat modeling, security testing, static and dynamic analysis, fuzz testing, vulnerability scans and penetration testing, manual code reviews, secure coding standards, security of open-source and third-party libraries, and protection of privacy and confidential data.

The model consists of five levels of maturity; none, initial, Basic, Acceptable, Mature. For each of the twenty parameters, five levels of maturity are defined, each with between 1–6 criteria that indicate whether a particular maturity level has been met for that practice. For instance, to achieve level 5 of the *Software Architecture and Design Reviews* parameter, you need to adhere to the following list of requirements:

1. Separation of privileges to address unknown attack vectors.
2. Reviews reveal multiple high and medium severity issues and the issues are effectively addressed early in the development cycle.
3. Architecture documents extensive enough to be used for Common Criteria (EAL-3) certification.
4. BSIMM-AA3.2: Drive analysis results into standard architecture patterns.

One of the more interesting parts of the PSMM is its inclusion of factors from other models (EAL-3, BSIMM-AA3.2) as adherence criteria. This leads to an explicit lists of requirements that the author would probably claim to be “the most suitable”, but also to some complexity in the model.

To perform a PSMM assessment, an organization first defines the scope of the assessment, which includes determining the products or systems that will be evaluated and the level of detail of the assessment. Next, key stakeholders are identified and involved in the assessment, as they are able to provide valuable insights and perspectives on the organization’s product security practices.

After the scope and stakeholders have been defined, the organization then collects and analyzes data on its product security practices. This involves reviewing documentation, conducting interviews, and gathering data from systems and tools. The data is then used to determine the organization’s current level of product security maturity, as well as any areas for improvement.

3 Research Approach

Object of Study. The study focuses on PSMM. The model was developed by Intel and is being used by a number of large IT companies including McAfee, Intel, and Deloitte. PSMM aims to be a simple, quantitative tool with low overhead that allows organizations to determine how well each Security Development Lifecycle activity is being performed. The PSMM is unique in that it provides relatively low-touch assessments, compared to more extensive models.

To perform this task, the model has operational parameters, such as Resources, Processes and Training, and technical parameters such as threat modelling and dynamic analysis. For each parameter five maturity levels are defined. Each of the maturity levels is associated with several questions per parameter.

If the answer to each of those questions is positive, the maturity level can be seen as obtained for that maturity level. As the model is simple and these levels are quantified and fully defined, minimal training and effort is needed to apply the model and create insightful metrics.

Evaluating Design Science Artifacts. Design science is the science of designing new information systems artifacts, that have a positive effect on science or society [12]. An essential step in the scientific process of design science, is the evaluation of design science artifacts. We frame our evaluation of the PSMM using Venable et al.’s framework [40]. The framework takes input from contextual factors such as goals, conditions, and constraints and supports the researcher in selecting the appropriate evaluatory techniques. These techniques are sorted into four categories that consist of two properties, being *ex post* (after creation of the artefact) or *ex ante* (before creation of the artefact) and a naturalistic (for example, in a field setting) or artificial (for example, in a laboratory) evaluation. After selecting one or more categories the framework proposes methods that can best be used with the selected evaluatory techniques.

Following the Design Science Research Evaluation Framework results in a focus on utility and efficacy. Essentially, posing that the evaluation should focus on the questions, ‘Does the model do what it needs to do?’ and ‘Can PSMM be effective?’. The framework subsequently suggests, based on contextual factors, that a naturalistic *ex post* approach is the best fit for this study. For this approach a number of methods are recommended including focus groups, surveys, and case studies. In this work, we use the case study method [32] for the evaluation, by performing a holistic multiple case study in Sect. 5.

4 Related Models

In this study, Snowballing was applied as the primary method to investigate the existing literature regarding the security maturity models. During the initial hypothesis search phase, we explored literature based on the following search keywords: “(security or SDL) maturity model”, and “Secure Development Lifecycle”. Accordingly, We collected a set of papers based on the snowballing method during this phase. Hence, we found 97 papers for security maturity models with different activities and features. Inclusion and exclusion criteria ensure that relevant manuscripts are included and irrelevant manuscripts are excluded. We extracted the required information, including the title, abstract, the Maturity Models considered in the paper, the venue where the paper was presented, the number of citations, and the year as inclusion and exclusion criteria.

The first and second authors conducted a quality assessment of the resulting studies. We collaboratively analyzed and discussed the studies for inclusion in the final list. We used quality criteria such as whether the paper contains (1) a problem statement, (2) research questions, (3) research challenges, (4) explicit research results, and (5) real-world use cases. Based on these qualities, we indicated each paper’s relevance to our study’s research question. Based on

this information, we have ranked the studies using four qualitative values: No relevance, low, medium, and high. The high-ranked results are listed in Table 1.

We ended up selecting 29 studies from various domains through a literature review based on snowballing that was presented in Table 1. We discovered that the studies we examined incorporated various security maturity models, such as BSIMM, SAMM, SSE-CMM, C2M2, MSSDL, CLASP, SAFECode, and Open-SAMM. However, upon analyzing the frequency of each framework's appearance in these studies, it became evident that BSIMM and SAMM were the popular choices. These two models demonstrated a consistent presence across the studies we considered in our research and they are open community projects and widely utilized within the IT industry.

OWASP Software Assurance Maturity Model (SAMM) - SAMM [35] is an open framework developed by OWASP, designed to assist organizations in assessing their current software security practices across the entire organization. This flexible model is intended for use by companies of all sizes, including small, medium, and large enterprises. SAMM is structured around key business functions within the software development life cycle, with each business function associated with three specific security practices. These business functions include Governance, Construction, Verification, and Operations [43].

Building Security In Maturity Model - BSIMM is founded on real-world practices observed in a large number of companies, making it a reflection of the prevailing state of software security. This framework is instrumental in evaluating the effectiveness of the Secure Software Development Lifecycle (SSDL). BSIMM covers 12 practices, which are further categorized into four primary domains: Governance, Intelligence, SSDL Touchpoints, and Deployment [16,19].

The practices and activities outlined in these models differ slightly in their approaches to what each model takes to achieve a higher maturity level. For instance, SAMM provides a comprehensive view by detailing activities, performance metrics, associated assurance benefits, personnel roles, and cost considerations. Conversely, BSIMM primarily focuses on security activities, the individuals engaged in them, and performance measurement [26].

We conducted a comparative analysis between PSMM and BSIMM, and SAMM. The results of this analysis are presented in the Table 2. The mappings were established based on comprehensive documentation and the respective activities defined in each model. In this mapping, we used a binary notation, with '1' denoting the presence of each activity from either the BSIMM or SAMM within specific parameters of the PSMM. For example, by considering the activity [SM1.1] from the "Strategy and Metrics" category, which involves 'publishing processes (roles, responsibilities, plan) and evolving them as necessary', we can realize that this particular activity can be effectively mapped to the "Process" parameter within the operational parameters of PSMM.

Table 1. An overview of the results of the literature study

Ref	Research type	Maturity Models
[21]	Research paper	BSIMM, SAMM, SSE-CMM, C2M2
[26]	Research paper	BSIMM, SAMM
[9]	Research paper	BSIMM, SAMM
[31]	Research paper	BSIMM, SAMM
[27]	Book	BSIMM, SAMM, MSSDL
[1]	Research paper	BSIMM, SAMM, MSSDL, CLASP, SAFECode
[22]	Research paper	BSIMM, SAMM, MSSDL, CLASP, SAFECode, OpenSAMM
[13]	Research paper	BSIMM, SAMM, MSSDL
[23]	Research paper	BSIMM, SAMM
[29]	Research paper	BSIMM, SAMM, MSSDL, CLASP
[3]	Research paper	BSIMM, SAMM, MSSDL
[30]	Research paper	BSIMM, SAMM, MSSDL, SSE-CMM
[34]	Research paper	BSIMM, SAMM, MSSDL, MSSDL, SAFECode
[44]	Research paper	BSIMM, SAMM, SAFECode
[20]	Research paper	BSIMM, SAMM
[18]	Research paper	BSIMM, SAMM
[37]	Thesis	BSIMM, SAMM, SAFECode
[41]	Research paper	BSIMM, MSSDL, CLASP, SAFECode
[45]	White paper	BSIMM, SAMM
[6]	Research paper	BSIMM, SAMM
[8]	Thesis	BSIMM, SAMM
[15]	Chapter	BSIMM, SAMM
[5]	Research paper	BSIMM, SAMM
[33]	Research paper	BSIMM, SAMM, MSSDL
[36]	Thesis	BSIMM, SAMM
[28]	Research paper	BSIMM, SAMM, MSSDL
[25]	Research paper	BSIMM, SAMM, MSSDL
[4]	Thesis	BSIMM, SAMM
[2]	Research paper	BSIMM, SAMM, MSSDL, CLASP, SAFECode

Through this mapping process, as shown in Table 2, we are able to quantify the number of activities from both BSIMM and SAMM that can be mapped to the PSMM framework. For activities where at least a '1' is assigned, it can be inferred that PSMM incorporates those activities within its scope. Thus, this analysis demonstrates the extent to which PSMM aligns with and covers activities outlined in BSIMM and SAMM. Moreover, in the coverage column, we indicated the activities and practices by '0' that they do not map to PSMM. For instance, the environment hardening practice in SAMM and part of the software environment practices in BSIMM. After analyzing this mapping, we realized that PSMM mapped to approximately 95% of the activities and practices outlined within BSIMM and it mapped to approximately 90% of the activities defined within SAMM (full table of mapping). On the other hand, PSMM assists organizations in advancing through the four stages of maturity management, establishing a clear path from their current product security status to the desired state. Within each stage of the maturity model, the team can showcase tangible achievements by evaluating specific requirements. This proactive approach outlined in the model enables the organization to set and reach milestones to

minimize product-related risks and detect potential risks earlier in SDL. The implementation of this maturity model will establish multiple layers of defense within the product, significantly raising the difficulty for malicious actors to breach it. The model's efficacy is evident at each security level as it enables the team to address security concerns in the early stages of development proactively.

5 Case Studies: 15 Software Producing Organizations

The case studies were performed at fifteen SPOs from 2021–2023. The organizations were companies ranging from one to 67,000 employees. In Table 3 the company sizes are indicated (Small: 1–49, Medium: 50–999, Large: 1000+). We do not provide exact numbers to protect the identity of some of the larger organizations, which are easily identifiable through their employee numbers. The PSMM was applied on one product per SPO. The organizations range from SPOs providing administration products for small businesses to SPOs producing products for maintaining public transportation vehicles. All SPOs are business to business companies. The SPOs are located in the Netherlands (12x), the USA (2x), and Canada (1x), although they all had a presence in the Netherlands. All interviews were conducted in Dutch and transcribed. The transcriptions are available upon request from the authors and were translated into English by the last author.

Case Study Protocol. The evaluation of the PSMM with experts was conducted by different student teams in the context of either a bachelor course at Utrecht University (Cases A-L) or in the context of a graduation project (M, N, O). A case study protocol ([Link to the case protocol](#)) was provided that included a case report format, a set of interview questions, and a guide to the PSMM. All teams were briefed in a two-hour session about the PSMM and about the case study approach in another lecture. Furthermore, they were provided with accompanying literature and prepared the case study interviews by discussing the protocol. All teams recorded their interviews and transcribed them. The case study data and PSMM assessment, collected by the researchers, consisted of: a filled in PSMM spreadsheet as provided by Toomey, spider graphs presenting the scores, a descriptive case study report (15–35 pages LNCS, available by request from the last author), and a transcription of the interviews performed (usually one or two per case study). The teams also reported on which document resources (website, provided documents, etc.) were used for the data gathering.

To analyze the effect of a company's size on the Operational, Technical, and combined scores, we use the Kruskal-Wallis (KW) test as our data are ordinal in nature and have more than two levels (small, medium, and large sizes). To explore any statistically significant results identified by the KW test, we use a post-hoc Mann-Whitney (MW) test (corrected for multiple tests with Bonferroni method). We adopt 5% as a threshold of α (i.e., the probability of committing Type-I error). We also provide the Cliff's δ , a non-parametric effect size measure, when reporting any statistically significant result identified with the MW test.

Table 2. The first table provides an overview of how PSMM maps to BSIMM, and the second table presents an overview of the mapping between SAMM and BSIMM. In this mapping process, we utilized a binary notation, where ‘1’ signifies the existence of each activity from either the BSIMM or SAMM within the defined parameters of the PSMM. For instance, examining the activity [CP1.3] in the ”Compliance & Policy (CP)” category of ”BSIMM” reveals that this specific activity can be effectively mapped to the ”Policy” parameter within the operational framework of PSMM. The full table for mapping PSMM - BISIMM and PSMM- SAMM is available as a spreadsheet at [this Google Drive Spreadsheet](#).

Activities (BSIMM - PSMM)		Operational Parameters										Technical Parameters							Coverage				
		Program	Resources	SDL	PSIRT	Policy	Process	Training	Reporting/Tracking	Sec. req. plan, DoD	Design reviews	Threat Modeling	Security Testing	Static Analysis	Dynamic Analysis	Fuzz Testing	Vuln and pen scans	Manual Code Reviews		Secure Coding Standards	Software supply chain	Privacy	
Governance																				88.24%			
STRATEGY & METRICS (SM)	LEVEL 1	[SM1.1]							1													1	
		[SM1.2]		1																			1
		[SM1.3]	1																				1
		[SM1.4]			1																		1
	LEVEL 2	[SM2.1]																					0
		[SM2.2]																					0
LEVEL 3	[SM2.3]	1																				1	
	[SM2.5]								1													1	
	[SM2.6]	1																				1	
COMPLIANCE & POLICY (CP)	LEVEL 1	[SM3.1]								1												1	
		[SM3.2]									1											0	
		[CP1.1]																					0
	LEVEL 2	[CP1.2]																			1	1	1
		[CP1.3]						1														1	1
		[CP2.1]																			1	1	1
[CP2.2]		1																			1	1	
LEVEL 3	[CP2.3]																				1	1	
	[CP2.4]																				1	1	
	[CP2.5]																				1	1	
[CP3.1]																					1		
[CP3.2]																				1	1		
[CP3.3]	1																				1		

Activities (SAMM - PSMM)		Operational Parameters										Technical Parameters							Coverage			
		Program	Resources	SDL	PSIRT	Policy	Process	Training	Reporting/Tracking	Sec. req. plan, DoD	Design reviews	Threat Modeling	Security Testing	Static Analysis	Dynamic Analysis	Fuzz Testing	Vuln and pen scans	Manual Code Reviews		Secure Coding Standards	Software supply chain	Privacy
Governance																				100.00%		
Strategy & Metrics	SM1-A	1																				1
	SM1-B	1																				1
	SM2-A	1																				1
	SM2-B	1																				1
	SM3-A										1											1
	SM3-B										1											1
Policy & Compliance	PC1-A					1																1
	PC1-B					1																1
	PC2-A					1														1		1
	PC2-B					1																1
	PC3-A						1															1
	PC3-B							1														1
Education & Guidance	EG1-A									1												1
	EG1-B									1												1
	EG2-A									1												1
	EG2-B										1											1
	EG3-A		1																			1
	EG3-B										1											1

The KW test identified statistically significant effect of the company’s size on the Operational and combined PSMM score ($p = 0.009$ and $p = 0.03$, correspondingly). For the Technical score the KW test returned $p = 0.15$ indicating no significant effect. The MW test requires the homogeneity of variance of samples.

Table 3. The 15 companies are listed here with their evaluation scores. The PSMM discriminates well across different companies, as many different values are given for different cases. The patterns in this table are discussed in Sect. 6.

	Company	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O		
	Size	M	S	S	L	L	M	L	S	S	S	M	M	S	S	S		
Operational Parameters																	Avg	StDv
O1	Program	5	4	1	5	5	5	4	1	1	2	2	4	3	1	1	4.17	1.71
O2	Resources	4	4	1	4	2	5	4	2	1	1	2	2	3	2	3	3.33	1.29
O3	SDL	1	3	3	3	5	4	3	0	1	1	5	5	3	2	1	3.17	1.63
O4	PSIRT	3	3	2	4	2	2	5	2	1	4	1	4	4	2	2	2.67	1.22
O5	Policy	3	4	4	4	3	5	3	1	1	5	4	4	3	2	1	3.83	1.36
O6	Process	1	5	2	5	5	4	5	4	3	4	2	4	4	1	3	3.67	1.41
O7	Training	2	2	1	5	4	5	5	1	1	3	4	3	2	0	3	3.17	1.62
O8	Reporting & Tracking	4	3	4	5	4	5	2	2	4	2	3	3	4	2	1	4.17	1.21
Technical Parameters																	Avg	StDv
T1	Sec. req. plan, DoD	1	5	2	5	4	5	4	2	3	2	0	4	4	2	2	3.67	1.56
T2	Design reviews	4	5	2	5	4	3	0	2	3	2	1	4	3	2	3	3.83	1.41
T3	Threat Modeling	2	1	2	4	2	4	3	3	3	1	0	3	1	1	1	2.50	1.22
T4	Security Testing	3	4	2	5	5	4	5	2	2	4	1	5	3	1	3	3.83	1.44
T5	Static Analysis	5	5	3	4	5	5	0	2	1	4	2	3	3	3	3	4.50	1.52
T6	Dynamic Analysis	4	4	0	4	5	5	0	1	0	4	2	2	3	2	2	3.67	1.77
T7	Fuzz Testing	1	4	0	5	5	4	3	0	1	1	1	3	1	1	1	3.17	1.75
T8	Vuln and pen scans	3	4	2	5	5	4	3	1	1	4	1	4	3	1	1	3.83	1.52
T9	Manual Code Reviews	5	3	4	5	3	4	4	3	4	5	3	5	1	2	3	4.00	1.18
T10	Secure Coding	2	3	3	5	3	4	3	3	3	3	1	3	4	5	1	3.33	1.16
T11	Software supply chain	2	4	3	4	3	5	4	1	1	2	1	4	4	4	0	3.50	1.52
T12	Privacy	3	4	2	5	4	0	3	4	4	4	3	3	4	2	1	3.00	1.33
	Operational score	2.9	3.5	2.3	4.4	3.8	4.4	3.9	1.6	1.6	2.8	2.9	3.6	3.3	1.5	1.9		
	Technical score	2.9	3.8	2.1	4.7	4.0	3.9	2.7	2.0	2.2	3.0	1.3	3.6	2.8	2.2	1.8		
	PSMM Score	2.9	3.7	2.2	4.5	3.9	4.1	3.3	1.8	1.9	2.9	2.1	3.6	3.0	1.8	1.8		

We checked this parameter with the Levene’s test confirmed that the samples for the three scores met this requirements (Levene’s $p > 0.61$). The post-hoc MW test with Bonferroni correction ($\alpha = 0.05/3 = 0.0167$) revealed several statistically significant results. For the Operational score we observed a statistically significant difference between Medium over Small (mean $Op_{small} = 2.16$ and $Op_{med} = 3.4$, MW $p = 0.014$ and Cliff’s $\delta = 0.83$, considered a large effect size) and Large over Small organizations (mean $Op_{small} = 2.16$ and $Op_{large} = 4.0$, MW $p = 0.0167$ and Cliff’s $\delta = 1$, large effect size). For the combined PSSM score the post-hoc test revealed similar trend between Small and Medium (mean $Op_{small} = 2.3$ and $Op_{med} = 3.16$, MW $p = 0.07$) and Small and Large organizations (mean $Op_{small} = 2.3$ and $Op_{large} = 3.89$, MW $p = 0.03$), but these results are not statistically significant.

We can draw several conclusions from the relationship between company size and PSMM score. First, the operational security within an SPO is directly related to its size. Second, technical security is not observably related to its size,

which can be explained by technical prowess: each company will have its own security requirements for a product and its skill levels, independent of size [14].

6 Analysis: Evaluating the PSMM

We evaluated the model in a free format; throughout interviews, the case study participants were allowed and encouraged to criticize parts of the PSMM during the assessment. At the end of the interviews, we also asked them what their general feelings about the model was. We report on these using quotes from the interviews and mark the finding with the companies where it was observed (e.g., A, B, C). If one of the companies' code names is in italics, that means the transcript shows this quote literally (company C in the example).

There were many positive remarks about the model. All organizations indicated that *“it is a great standardized test to benchmark one’s operational security”*. While we never shared the data from other organizations with them, the benchmarking capabilities were still recognized. Another positive remark we heard from the participants concerned that it was timely to take a look through this lens. Each organization found low hanging fruits for improvement, and this generally helped the organization. A final positive remark we heard was about how to prioritize security in the software development process: *“The model proved useful to us, because we typically prioritize features over security, we should start writing security “features” down as user stories”* (H, I, K).

We collected 24 unique criticisms from the interviews, after grouping them for occurrence. The following texts report on the ones that are common (three or more companies) or stand out for other reasons.

Completeness - The participants were particularly critical of the model completeness. Most of them found it *“overcomplete”* (F, G, L, K, M, N, O) and *“practically impossible to be fully compliant”* (K, M, N, O) *“without huge budgets”* (all). For example, one participant mentioned that if you follow the model strictly *“being available 24/7 is a requirement, so maximum maturity cannot be reached, because we don’t need 24/7 availability”* (F). On the other hand, it was judged to be *“more or less sufficient for what it’s trying to do”* (A, F, D).

Flexibility - *“Maturity Models are generally too static”* (A, B, L, K), and the participants want the *“Model [to] be more ‘need-based’, and take the company goals into account.”* (F, K). Furthermore, the PSMM is judged to be *“too strict on particular guidelines, e.g. ISO”* (A, B, D, G, J, K, M, N, O)

Score Representation and Correctness - One important critique was also that the comprised score that is assigned at the end of the process does not fairly represent the status of a company and can be *“misleading”* (A, D, K, M, N, O). A relevant detail is that the way in which the score is calculated in the provided spreadsheets, is different from how it is described in the description text of the

model. Some organizations also wondered whether the model might give “*a false sense of security*” (A, F, D).

Security Culture - Some of the case participants that found the model too inflexible, also mentioned that the model insufficiently allows for situationality in security culture. This was observed on different levels, such as culture on the work floor: “*The model assumes zero trust within the company itself, which may be an American thing.*” (A, E, L), but also the situation that customers of a product may be more demanding regarding security and may be more vigilant and in a more trusting relationship with the SPO.

Assessment Complexities - One interesting complexity was that in some of the cases, we could not find all details on security processes, as they had “*some processes ... outsourced, such as pen testing*” (C, G, L). Furthermore, we heard from some organizations that by “following modern certifications for security, we scored high by default” (E, F). In larger organizations, we also encountered case participants who did not precisely know how particular functions were filled in within the organization (E).

6.1 PSMM Usability and Situational Factors

The PSMM instructions are somewhat unclear on its use; should the PSMM be applied regularly or is it a one-time instrument? Should the scores be trusted and have an impact on the improvement policies within the organization? And for whom is the model suitable? In this Section, we answer those questions using the evaluations and general knowledge about maturity models.

The models are generally tailored towards larger organizations, and the PSMM, with its origins at Intel, seems to suffer from this more than others. This has some funny side effects, such as interpretations leading to smaller (single product) organizations being able to much more rapidly adhere to some of the requirements. For example, to achieve level 5, an organization needs to have a Product Security Champion for a product, which is relatively easy for a one-product company.

For some of the other requirements, the inverse is true. A small-scale organization would not be able to meet some of the other requirements or only with immense and unnecessary difficulty. An example of this can be found in the resources parameter; To achieve level three the organisation needs to have a budget for the growth of the number of product security champions and have one product security champion per product. However, if a small organization has only a single product with a product security champion, then budgeting for multiple new product security champions seems unnecessary.

Situational Factors. A situational factor is any factor relevant to product development and product services. Examples are company size, branch and

the number of submitted requirements per month, whether or not currently a waterfall-based method is used for product software development, etc. [7]. The organization's context is considered by evaluating different situational factors that define its surroundings and structure, subsequently helping the choice of relevant capabilities [7]. We suggest incorporating two situational factors that could improve the PSMM. Such factors can serve multiple purposes: they can either automatically disregard or introduce specific practices, or they can facilitate branching within the model to another variation. After identifying four potential situational factors through the interviews, we have chosen to introduce only two of them as real options.

The first situational factor we identify is company size. There are two sides of the spectrum that the interviewees addressed: small one-product companies should be given exemptions from practices in the model. On the other hand, large organizations require flexibility for the implementation of processes, as they may have more or less centralized security services within the organization, and at times the PSMM is too prescriptive in this respect. The second situational factor we identify is “*the development method (agile or waterfall)*” (A, H, I), especially because agile takes a different approach to security [30].

There were also proposed situational factors that we mention here, but question the validity of, and we currently do not propose implementing them in the PSMM. The third situational factor concerns the product characteristics, with two variation points. First, one of the companies operates from an open source perspective and provides a large part of its code base to the open source community (D), inherently leading to more secure products. One of the participants stated that product maturity has a strong influence on security; “it’s easier to score better with a mature product.” (F, H, I, K).

General Usage and Frequency. From the case studies we find that the model is best usable for medium to large product organizations with multiple products. As future work, we propose that a lighter version of the model is developed for smaller one-product companies. Assessments can be done in a relatively short time, ranging from around four to eight hours to get a first score, but obviously the lessons are found in the next steps: where is the organization now, where does it want to go, and how does the PSMM help in deciding what to do next? With regards to maturity models [17, 24, 39], from experience we can say that a yearly assessment is frequent enough and many organizations only use the same maturity model for one to four iterations, after which they abandon the maturity model or move on to another more extensive model.

6.2 Threats to Validity

Conclusion Validity. Possible threats to conclusion validity are related to the inaccurate data and data analysis process. Each of the case study reports was checked by one of the authors using the associated transcript, which are available upon request from the last author. Furthermore, two lower quality case study

reports were excluded from the study, because they were incomplete and did not appear to represent the data. As for data analysis, we used the non-parametric tests as they do not require a normal distribution of the sample. To mitigate low statistical power, we adopted $\alpha = 0.05$ for the difference test, with reported Cliff's δ effect sizes for significant results.

Internal Validity. To perform the maturity assessments, we used the instructions as provided with the PSMM. We strongly depended on the information provided by the interviewees, and when vague answers were given, we were critical to ensure that we did not assess a practice or capability as present when it was not. The interviews had a dual nature: we performed the assessment and simultaneously asked the interviewee to provide feedback on the PSMM itself. This may have influenced the correctness of our findings, but we often found that asking deeper questions about each practice, led to better more detailed assessments and better shared understanding of each of the practices.

External Validity. To ensure the generalisability of our findings, we conducted a series of case studies with real product companies of different sizes, backgrounds, and from different regions. Therefore, we collected a diverse set of cases of applying the PSMM to evaluate the security maturity of real product development cases. However, it should be noted that we refrain from making any claims to generalization, but that we suspect that the PSMM is suitable for use by medium SPOs. We find that our model observations in this Section are rather generic and could be made about other maturity models or security assessment models as well. We hope that in the future, model designers will take these challenges into account, especially regarding applicability and situationality.

7 Conclusion

In this work, we provide an academic evaluation of a model rooted in practice entitled the Product Security Maturity Model, by evaluating it with 15 case studies and comparing it to existing models. We provide an extensive criticism of the model itself and how it may be improved, but we also praise it for its usefulness and effectiveness in providing organizations with improvement advice. We identify several situational factors that could lead to variations in the model that better fit an organization's size or development method.

We observe that maturity models are a well accepted standard for the diffusion of knowledge in organizations and are frequently used within organizations with highly skilled workers, such as in information technology. The 15 case participants all agree that even though the model is not perfect, it immediately gave the interviewees new ideas and concepts to implement and check within the organization. As such, we dare state that our work has already made an impact at the time of writing this work.

As part of our future work, we consider exploring other models and their applicability to software businesses, also to circumvent the challenges that have

been identified in Sect. 6. In December 2023 we will start a new set of case studies with the OWASP SAMM 2.0 model. We experience that maturity models are seen as a relevant instrument for disseminating (scientific) knowledge among organizations, but are not necessarily seen as scientific. After all, aren't they just collections of ideas without much scientific merit? We consider it a challenge to give maturity models more solid footing in the scientific community, for instance by performing more empirical studies on the longevity of maturity models and their usage. We have already created a platform for the dissemination of maturity models and ensure their visibility: [MaturityModels.org](https://maturitymodels.org).

Acknowledgments. We want to thank the student teams that so diligently performed the case studies according to our protocol.

References

1. Al-Matouq, H., Mahmood, S., Alshayeb, M., Niazi, M.: A maturity model for secure software design: a multivocal study. *IEEE Access* **8**, 215758–215776 (2020)
2. M. Alenezi, H. A. Basit, M. A. Beg, and M. S. Shaukat. Synthesizing secure software development activities for linear and agile lifecycle models. *Softw.: Pract. Exp.* **52**(6), 1426–1453 (2022)
3. Ardo, A.A., Bass, J.M., Gaber, T.: An empirical investigation of agile information systems development for cybersecurity. In: Themistocleous, M., Papadaki, M. (eds.) *Information Systems: 18th European, Mediterranean, and Middle Eastern Conference, EMCIS 2021, Virtual Event, December 8–9, 2021, Proceedings*, pp. 567–581. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-030-95947-0_40
4. Assal, H.: The human dimension of software security and factors affecting security processes. PhD thesis, Carleton University (2018)
5. Assal, H., Chiasson, S.: Security in the software development lifecycle. In: 14th Symposium on Usable Privacy and Security (SOUPS 2018), pp. 281–296 (2018)
6. Attwood, S., Onumah, N., Paxton-Fear, K., Kharel, R.: Security-focused prototyping: A natural precursor to secure development. In: 2022 13th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), pp. 356–361. IEEE (2022)
7. Bekkers, W., Spruit, M.R., van de Weerd, I., van Vliet, R. and Mahieu, A., et al.: A situational assessment method for software product management. In: *Proceedings of the 18th European Conference on Information Systems (ECIS2010)* (2010)
8. Bideh, P.N.: Contributions to Securing Software Updates in IoT. Department of Electrical and Information Technology, Faculty of Engineering (2022)
9. Bugeja, J., Vogel, B., Jacobsson, A., Varshney, R.: IoTSM: an end-to-end security model for IoT ecosystems. In: 2019 International Conference on Pervasive Computing and Communications Workshops, pp. 267–272. IEEE (2019)
10. Farshidi, S.: Multi-criteria decision-making in software production. PhD thesis, Utrecht University (2020)
11. Hathaway, O.A., et al.: The law of cyber-attack. *California law review*, pp. 817–885 (2012)
12. Hevner, A., Chatterjee, S., Hevner, A., Chatterjee, S.: Design science research in information systems. *Design research in information systems*, pp. 9–22 (2010)

13. Höst, M., Hell, M.: Evaluation of the havoss software process maturity model. In: 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 137–140. IEEE (2020)
14. Hou, F., Jansen, S.: A systematic literature review on trust in the software ecosystem. *Empir. Softw. Eng.* **28**(1), 8 (2023)
15. Iovan, M., Cruzes, D.S., Johansen, E.A.: A framework for a sustainable software security program. *Evolving Software Processes*, pp. 47–69 (2022)
16. Jaatun, M.G.: The building security in maturity model as a research tool. In: *Empirical Research for Software Security*, pp. 201–208. CRC Press (2017)
17. Jansen, S.: A focus area maturity model for software ecosystem governance. *Inform. Softw. Technol.* **1**, 118 (2020)
18. Kudriavtseva, A., Gadyatskaya, O.: Secure software development methodologies: a multivocal literature review. arXiv preprint [arXiv:2211.16987](https://arxiv.org/abs/2211.16987) (2022)
19. McGraw, G.: Software security and the building security in maturity model (bsimm). *J. Comput. Sci. Coll.* **30**(3), 7–8 (2015)
20. Moyón, F., Bayr, C., Mendez, D., Dännart, S., Beckers, K.: A light-weight tool for the self-assessment of security compliance in software development – an industry case. In: Chatzigeorgiou, A., Dondi, R., Herodotou, H., Kapoutsis, C., Manolopoulos, Y., Papadopoulos, G.A., Sikora, F. (eds.) *SOFSEM 2020: Theory and Practice of Computer Science: 46th International Conference on Current Trends in Theory and Practice of Informatics, SOFSEM 2020, Limassol, Cyprus, January 20–24, 2020, Proceedings*, pp. 403–416. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-38919-2_33
21. Nikbakht Bideh, P., Höst, M., Hell, M.: HAVOSS: a maturity model for handling vulnerabilities in third party OSS components. In: Kuhrmann, M., Schneider, K., Pfahl, D., Amasaki, S., Ciolkowski, M., Hebig, R., Tell, P., Klünder, J., Küpper, S. (eds.) *PROFES 2018. LNCS*, vol. 11271, pp. 81–97. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03673-7_6
22. Núñez, J.C.S., Lindo, A.C., Rodríguez, P.G.: A preventive secure software development model for a software factory: a case study. *IEEE Access*, **8**, 77653–77655 (2020)
23. Onumah, N., Attwood, S., Kharel, R.: Towards secure application development: A cyber security centred holistic approach. In: 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), pp. 1–6. IEEE (2020)
24. Overeem, M., Mathijssen, M., Jansen, S.: Api-m-famm: a focus area maturity model for API management. *Inform. Software Tech.* **147**, 106890 (2022)
25. Palma, F., Realista, N., Serrão, C., Nunes, L., Oliveira, J., Almeida, A.: Automated security testing of android applications for secure mobile development. In: 2020 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), pp. 222–231. IEEE (2020)
26. Ramirez, A., Aiello, A., Lincke, S.J.: A survey and comparison of secure software development standards. In: 2020 13th CMI Conference on Cybersecurity and Privacy, pp. 1–6. IEEE (2020)
27. Ransome, J., Misra, A.: *Core software security*. CRC Press (2018)
28. Rindell, K., Holvitie, J.: Security risk assessment and management as technical debt. In: 2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), pp. 1–8. IEEE (2019)
29. Rindell, K., Hyrynsalmi, S., Leppänen, V.: Aligning security objectives with agile software development. In: *Proceedings of the 19th International Conference on Agile Software Development: Companion*, pp. 1–9 (2018)

30. Rindell, K., Ruohonen, J., Holvitie, J., Hyrynsalmi, S., Leppänen, V.: Security in agile software development: a practitioner survey. *Inf. Softw. Technol.* **131**, 106488 (2021)
31. Rindell, K., Ruohonen, J., Hyrynsalmi, S.: Surveying secure software development practices in finland. In: *Proceedings of the 13th International Conference on Availability, Reliability and Security*, pp. 1–7 (2018)
32. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empir. Softw. Eng.* **14**, 131–164 (2009)
33. Ryan, I., Roedig, U., Stol, K.-J.: Insecure software on a fragmenting internet. In: *2022 Cyber Research Conference-Ireland (Cyber-RCI)*, pp. 1–9. IEEE (2022)
34. Ryan, I., Roedig, U., Stol, K.-J.: Measuring secure coding practice and culture: A finger pointing at the moon is not the moon. In *2023 IEEE/ACM 45th Int'l Conference on Software Engineering (ICSE)*, pp. 1622–1634. IEEE (2023)
35. Teodoro, N., Serrão, C.: Web application security: improving critical web-based applications quality through in-depth security analysis. In: *International Conference on Information Society (i-Society 2011)*, pp. 457–462 (2011)
36. Tøndel, I.A.: Prioritisation of security in agile soft. dev. projects (2022)
37. van de Werfhorst, M., Poll, E., Schoemaker, H.: and C. Kop, Security recommendations for agile and devops development at ridder data systems (2020)
38. van Steenbergen, M., Bos, R., Brinkkemper, S., van de Weerd, I., Bekkers, W.: The design of focus area maturity models. In: Winter, R., Zhao, J.L., Aier, S. (eds.) *Global Perspectives on Design Science Research*, pp. 317–332. Springer Berlin Heidelberg, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13335-0_22
39. van Steenbergen, M., Bos, R., Brinkkemper, S., van de Weerd, I., Bekkers, W.: Improving is functions step by step: the use of focus area maturity models. *Scand. J. Inf. Syst.* **25**(2), 35–56 (2013)
40. Venable, J., Pries-Heje, J., Baskerville, R.: Feds: a framework for evaluation in design science research. *Eur. J. Inf. Syst.* **25**, 77–89 (2016)
41. Venson, E., Alfayez, R., Gomes, M.M., Figueiredo, R.M., Boehm, B.: The impact of software security practices on development effort: An initial survey. In: *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 1–12. IEEE (2019)
42. Von Solms, R., Van Niekerk, J.: From information security to cyber security. *Comput. Secur.* **38**, 97–102 (2013)
43. Wen, S.-F.: Software security in open source development: a systematic literature review. In: *2017 21st Conference of Open Innovations*, pp. 364–373. IEEE (2017)
44. White, C.A.: Root causes of insecure internet of things and holistically addressing them. In: *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1066–1074. IEEE (2020)
45. Williams, L.: Secure software lifecycle knowledge area issue. The National Cyber Security Center (2019)



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Strategic Digital Product Management in the Age of AI

Helena Holmström Olsson¹ and Jan Bosch²

¹ Malmö University, Malmö, Sweden
helena.holmstrom.olsson@mau.se

² Chalmers University of Technology, Göteborg, Sweden
jan.bosch@chalmers.se

Abstract. The role of software product management is key for building, implementing and managing software products. However, although there is prominent research on software product management (SPM) there are few studies that explore how this role is rapidly changing due to digitalization and digital transformation of the software-intensive industry. In this paper, we study how key trends such as DevOps, data and artificial intelligence (AI), and the emergence of digital ecosystems are rapidly changing current SPM practices. Whereas earlier, product management was concerned with predicting the outcome of development efforts and prioritizing requirements based on these predictions, digital technologies require a shift towards experimental ways-of-working and hypotheses to be tested. To support this change, and to provide guidelines for future SPM practices, we first identify the key challenges that software-intensive embedded systems companies experience with regards to current SPM practices. Second, we present an empirically derived framework for strategic digital product management (SPM4AI) in which we outline what we believe are key practices for SPM in the age of AI.

Keywords: Strategic digital product management · DevOps · Data · Artificial intelligence · Digital ecosystems · Digitalization · Digital transformation

1 Introduction

The role of product management is critical for the success of any product. As recognized in [7], the product manager holds responsible for product requirements, release definition, product release lifecycles, creating an effective product introduction team and preparing and implementing the business case. Similarly, [27] describes software product management (SPM) as a crucial discipline that encompasses the activities and responsibilities involved in creating, delivering, and maintaining software products. In addition, and as pointed out in [7], the product manager owns the business case and assures that a product release delivers the expected value to customers as well as to the business. In practice,

and especially in the software-intensive embedded systems industry, SAFe is one of the most common frameworks for product strategy, planning and roadmapping¹. During recent years, it has become widely adopted by companies that wish to scale their agile practices, accelerate value delivery and shorten feedback loops to customers. Although research on the benefits of adopting SAFe is still scarce, it remains the predominant framework for software organizations that seek to accelerate value delivery to customers. In addition to SAFe, there are several frameworks and models for supporting and improving software product management practices. As a few examples, the ISPMA framework provides a holistic view on the activities of software product management², the SPM reference framework identifies key process areas as well as the stakeholders and their relations [35], the SPM competence model outlines key capabilities a software organization should implement to improve SPM maturity [2], the market-driven product management and requirements engineering model (MDREPM) enables software process improvement and process assurance [13] and the 4CC model provides a blueprint for re-engineering product development management practices [30]. Also, there are numerous papers outlining key success factors for software product management [8] and SPM best practices, e.g., [10,33,36].

However, although there is prominent research on software product management and the importance of this discipline, there are few studies that explore how the role of product management is rapidly changing due to recent, and profound, trends that come with digitalization and digital transformation. As concluded in our previous research [5], digital technologies change development organizations and how these operate. In our view, digital transformation has significant implications on the software product management. Similarly, [21] recognize how the principles of how software products are introduced and delivered to customers are changing rapidly. Although software product management can, and in our view should, be considered part of the field of software engineering, in this paper we use these terms as separate. In the remainder of the paper, we use the term software product management to refer to decisions concerning what to build and why it should be built. We use the term software engineering to refer to decisions and activities concerning how to build the prioritized functionality.

In this paper, we explore how key trends such as DevOps, data and artificial intelligence (AI), and the emergence of digital ecosystems challenge and fundamentally change current SPM practices. Our research builds on multi-case study research in companies in the embedded systems domain that experience rapid changes in the business environments in which they operate and as a consequence, need guidelines for how to approach and reason about their SPM practices going forward.

The contribution of this paper is two-fold. First, we identify the key challenges that companies in the software-intensive embedded systems domain experience with regards to their current SPM practices. Second, we present an empirically derived framework for strategic digital product management (SPM4AI) in which

¹ <https://scaledagileframework.com/>.

² ispma.org.

we outline what we believe are key practices for SPM in the age of artificial intelligence (AI).

The remainder of this paper is structured as follows. In Sect. 2, we review literature on software product management and frameworks that are currently used to support this role. Also, we outline key trends that we see challenge current SPM practices. In Sect. 3, we provide an overview of the research approach we used and the case companies involved in our study. In Sect. 4, we present the empirical findings. In Sect. 5, we present the 'Strategic digital Product Management' framework (SPM4AI) in which we outline what we believe are key practices for SPM in the age of AI. In Sect. 6, we discuss threats to validity. In Sect. 7, we conclude the paper.

2 Background and Related Work

2.1 Software Product Management (SPM)

Engineering is concerned with building systems and with activities such as e.g., requirements engineering, designing an architecture, developing software, implementation of software, testing and validation of the system and finally, release to customers. However, whereas engineering is concerned with 'how' to build systems, there is another activity concerned with 'what' to build and even more important, 'why' we should build the system in the first place. This activity is typically referred to as product management and in the context of software-intensive systems as software product management. Over the years, numerous studies have explored the activities involved in software product management and the role of the software product manager. In [8], the authors conclude that the SPM role is critical and that with a consistent and empowered product management role, the success rate of projects in terms of schedule, predictability, quality and project duration improves. In [2], a product manager is referred to as the *"mini-CEO of an organization"* as they are positioned at the center of the organization where they keep in contact with all stakeholders to ensure that they work towards the same goal. In [28], the author discusses how proper product management processes improve resource management efficiency, lead to increased business growth, better budget control, higher user satisfaction, increased release predictability and faster release cycles. As depicted in [12], software product management is the role responsible for what the product is, how it works, whom it serves and how it affects the company and its customers. As a comprehensive summary, [32] outline key product management practices in a framework involving management processes, support processes and software lifecycle processes. As can be seen in the studies mentioned above, and if looking at the impressive body of knowledge in the field, the importance of this role is only increasing.

2.2 SPM Frameworks

There are several frameworks and models that provide support for software product management. With a focus on how to effectively scale agile practices,

SAFe has become one of the most common and widely adopted frameworks in industry [<https://scaledagileframework.com/>]. In the most recent version, product management is described as the function responsible for defining desirable, viable, feasible, and sustainable solutions that meet customer needs and as the function supporting development across the product life cycle. In [29], the authors conclude that increased transparency, alignment, quality, time to market, predictability and productivity are the perceived benefits of SAFe, while the challenges are associated with resistance to change and controversies with the framework.

In addition to SAFe, there are prominent frameworks such as e.g., the ISPMA framework [ispma.org]. This framework provides a holistic view on the activities of software product management with the intent to establish and improve SPM practices in organizations. In [15], the authors build on the ISPMA framework when providing best practices for product strategy, product planning, strategic management and orchestration of the functional units of the company. In [11], the framework is referred to as unique in that it integrates several key characteristics from previous frameworks for product management, as well as for student education purposes.

The SPM reference framework identifies key process areas as well as the stakeholders and their relations [35]. The framework is based on a review of state-of-the-art literature on software product management as well as experience from industrial case studies. In addition to this framework, the SPM competence model outlines key capabilities a software organization should implement to improve SPM maturity [2]. The model provides an overview of four business functions that are important to SPM, i.e., portfolio management, product planning, release planning and requirements management, and the focus areas for each of these functions. Also, the model indicates the interactions that take place between different stakeholders and how information flows between roles and functions.

As yet another model, the market-driven product management and requirements engineering model (MDREPM) enables software process improvement and process assurance in market-driven software engineering [13]. The model targets the unique challenges that product development organizations operating in market-driven environments are facing and can be seen as both a best-practice guide and a process assessment framework.

Finally, the 4CC (Four Cycles of Control) framework combines business management and software product development, and takes both a long-term and short-term view to software product release management [30]. The framework involves the type, timing, and content of different product releases, and aims at providing a common understanding for how to organize software product development.

2.3 Key Trends that Challenge Current SPM Practices

Based on recent research, as well as our experience of working closely with companies in the embedded systems domain for more than a decade, we identify

three trends that have an impact on current ways-of-working and that challenge current SPM practices. Below, we detail these trends and the effect they have on SPM.

DevOps: The emergence of agile practices was key as the sprint model fundamentally changed the ways in which software was developed and delivered. These practices are now scaling and with the emergence of DevOps the entire feedback cycle with customers is shortened when bringing development and operations together [18]. For DevOps to be effectively adopted, technical transformations include, e.g., automated deployments using build and continuous integration tools, treating infrastructure as code, and continuous monitoring of infrastructure and system behavior in production. On the organizational side, it is crucial to build and strengthen a collaborative culture to successfully establish a straightforward communication and shared responsibilities [9]. With DevOps, also the role of the product manager changes. First, it becomes much more integrated with the engineering team as the ways-of-working shift from being specification-centric to more experiment-centric. Second, with DevOps systems are grown instead of built. Rather than defining the requirements and building the system to meet the specification, the focus shifts to defining outcomes and iteratively deploying functionality that support these. Third, with an experiment-centric approach, product managers can continuously measure the impact of development efforts and hence, adopt a more customer-centric approach to product development.

Data and AI: Digital technologies are transforming industry to an extent that we have only seen the beginnings of. Across domains, companies experience rapid changes to their existing practices due to the many opportunities these technologies bring. As recognized in e.g., [5,25], data and AI allow for continuous improvement of system functionality and hence, continuous value delivery to customers. In addition, and as recognized in [26], data and AI provide the basis for new digital offerings and recurring revenue streams. Finally, data and AI enable companies to shift towards customer KPI-based business models and two-sided markets [1,31]. With data and AI, the role of the product manager shifts from being concerned with predicting the outcome of development efforts and prioritizing requirements based on these predictions, towards adopting experimental ways-of-working, defining hypotheses to be tested and using data from products in the field for continuous monitoring and improvement of customer value.

Digital Ecosystems: As a recent trend, business environments are being recognized as digital ecosystems [16]. The concept of digital ecosystems is proposed as a new way to perceive the increasingly complex and interdependent systems that are being created and that are characterized by self-organization, scalability, sustainability and with business models in which the main revenue stream no longer consists of the production of a product that is sold to customers, but rather, provision of a combination of services and products to their customers [16,17]. From a product management perspective, digital ecosystems reshape the business ecosystems in which companies operate. With new innovation platforms

and digital marketplaces, software product development is rapidly shifting from focusing on internal scale, efficiency, quality and serving customers in a one-to-one relationship, to contributing to an ecosystem of multiple players [4].

3 Research Method

3.1 Case Study Research

Case study research has become an appreciated method in software engineering research as it allows for empirical investigation of contemporary phenomena. In [3], case studies are defined as information gathering from a few selected entities with little or no experimental control. Similarly, [34] emphasizes how case studies are useful when studying organizational contexts with complex and intertwined conceptual structures. In our study, we adopted a multi-case study approach to explore how key trends such as DevOps, data and artificial intelligence (AI), and digital ecosystems challenge current SPM practices. The findings we present are based on close collaboration with a selected set of companies in the embedded systems domain. All the case companies are members of a larger research collaboration in which industry and academia work closely together to help accelerate digitalization (www.software-center.se). In what follows, we report on research in which we use company workshops and frequent check-in meetings conducted between January 2023 and September 2023 as the basis for our findings. It should be noted however, that we have been working with the case companies as part of the larger research initiative for more than a decade. This gives us the opportunity to use previous insights and experiences as valuable and complementary input also in this study.

3.2 Case Companies

The following case companies were involved in our study:

- *Case company A* is a networking and telecommunications company. For the purpose of this paper, we engaged with roles involved in software product management, engineering management and data analytics. We studied two different use cases. *Use case 1* is concerned with how to balance requests from a large and diverse customer group. *Use case 2* is concerned with how to effectively use data for continuous improvement of software products.
- *Case company B* is a company manufacturing vehicles. For the purpose of this paper, we engaged with roles involved in software product management, technology management and strategy lead. We studied one use case concerned with how to adopt A/B testing practices in large-scale systems development.
- *Case company C* is a food packaging and processing company. For the purpose of this paper, we engaged with roles responsible for data management and connectivity and software and systems engineering. We studied one use case concerned with using deep learning (DL) for managing system evolution.

- *Case company D* is a company manufacturing trucks. For the purpose of this paper, we engaged with roles responsible for product management, technology management and autonomous drive. We studied one use case concerned with using reinforcement learning to improve system behaviors.

3.3 Data Collection and Data Sources

As the primary data source for this study, we engaged in workshop sessions at all case companies. The workshop sessions lasted for 1–3 hours, involved 4–10 people and focused on current SPM practices, challenges imposed with digital technologies and best practices and strategies for how to address and mitigate these challenges. In addition to the workshops, we had bi-weekly and/or monthly check-in meetings to review status of the initiatives and we continuously discussed solution development and next steps. Our findings build on company workshops and frequent check-in meetings conducted between January 2023 and September 2023. We have worked with several of the case companies for more than a decade, and have reported on specific teams, products and challenges in previous work. However, in this paper the focus is on software product management whereas in earlier publications we focused on software engineering challenges. In total, we met with the case companies in 12 workshops (7 workshops in company A, 3 workshops in company C and 2 workshops in company D). With company B, we interacted primarily by using frequent check-in meetings (on-line) and e-mail conversations. The longitudinal nature of our research allows us to capture not only our most recent experiences in the companies, but also challenges and solutions that we have seen emerge over time as a result of their long-term and on-going digital transformation. As part of the collaboration with the case companies, we were able to follow several improvement initiatives as well as internal discussions on how to rethink and reinvent the SPM role.

4 Findings

The challenges experienced in the case companies are due to the rapid pace of digital transformation and the new technologies and ways-of-working that come with digitalization. From the perspective of SPM, this implies that existing frameworks are insufficient as these often fail in effectively supporting short DevOps cycles involving continuous development and delivery of data and AI-intensive system components. Below, we describe a selected set of use cases. Each use case illustrates a key challenge that the case company experience and how the company responded to this challenge.

4.1 Everything Starts with a Requirement

Challenge: The case companies develop systems that are safety-critical and subject to strict regulations and legislation. Due to this, the primary approach to

development in all case companies is a requirement driven approach. As the start of development, product management is responsible for collecting and specifying requirements as input for the software development teams. Over the years, and increasingly so with practices such as continuous deployment and data-driven development being introduced, a number of limitations have been recognized in relation to the requirement driven development approach. The assumption that customer requirements can be identified before development starts is the most questioned one and with an increasing amount of product and customer data available the traditional approach to requirements is rapidly changing.

Response: In the case companies, we notice that a requirement driven approach to development is well suited for situations in which features and functionality are well-understood, where there is a long-term agreement between the customer and the development organization and where there is less frequent change imposed on the system. However, when applied also in a fast changing environment the requirement driven approach falls short. This was confirmed in all case companies involved in our study and people report on use cases in which SPMs "create a false illusion of certainty" by taking a requirement driven approach also in situations characterized by uncertainty. Our research shows that a key challenge is to find alternative approaches and frameworks that support software product managers also in evolving and uncertain system contexts [6].

4.2 Balancing Exploration and Exploitation

Challenge: Case company A delivers systems to a large number of customers with very different needs. The role of product management is to inventory these needs, to combine, merge, and prioritize among them, and to present a roadmap with a set of requirements for the next release of the system. In this process, effective management of customer feedback is critical. However, and as reported in our previous work [23], the development of systems that serve a large customer group can easily lead to a tension between two conflicting interests. On one hand, the development organization seeks to achieve scale in terms of implementing as many new features to as many customers as possible. On the other hand, the development organization needs to show responsiveness to strategic customers. This requires the ability to balance exploration and exploitation which is a challenge in the companies we studied. In [23], we reported on the software engineering aspects of this by outlining the development organization and the structure of the software teams.

Response: From a SPM perspective, use case 1 in company A illustrates the challenge of balancing individual customer requests while at the same time serving a large customer. During the workshops in company A, we learnt that the most rewarding approach is to have some of the organization's development teams dedicated to specific customers that the product manager identifies as the most strategic ones. Based on the requests from these customers, teams explore new features, collect customer feedback and improve these features in an iterative

and incremental fashion. Once exploration of features is done with strategic customers, these features are adapted to generic customer needs and included in the planned releases. For the software product managers we talked to during this study, this approach allows for exploratory development of new features and the ability to respond rapidly to strategic customers, while over time having the benefit of exploiting these development efforts with the larger customer base. From the perspective of the software product managers involved in our study, the opportunities for exploration are rapidly increasing with large amounts of data, as well as AI technologies, being available.

4.3 Towards Testing of Hypotheses

Challenge: To manage situations with low certainty is a challenge that all case companies experience. Over the years, we studied cases where product management prioritized features that, in the end, were never used by customers or used so seldom that the development efforts could not be justified. To address this challenge, companies need support for experimental ways-of-working where teams use hypotheses instead of requirements as the basis for development as highlighted in data-driven development approaches. Although there is detailed advice for how to conduct A/B testing in online contexts, support for how and when to adopt these practices in large-scale embedded software development is scarce. Still, there are some examples from the companies we studied where experiments are run to support smaller improvements of features and where collection and analysis of customer and product data informs development.

Response: In company B, A/B testing is used on test vehicles with the intent to test two different versions of an energy optimisation software with customers. The test fleet consists of 28 vehicles and the company uses an experiment group design method, i.e., 'Balance Weight Matched Design', to address the challenge of having a limited sample size and increase the experiment power with small samples. In [19, 20], we present the software engineering aspects of these experiments and show that balanced groups can be produced even when the sample sizes are small. Our recent interactions with product managers in case company B confirm that experimentation is well suited for situations where there is a need to test different hypotheses and where the solution to a problem is unclear. Also, the company has started applying experimentation in innovation efforts as there is the need to test and trial with customers in order to identify the potential value of new digital services and offerings.

4.4 Maximizing Use of Big Data Sets

Challenge: The case companies collect massive amounts of data from their products in the field. This data is primarily used for diagnostics and quality assurance as well as for monitoring and improving product performance. Most companies experience a situation in which the amounts of data are growing exponentially

due to an increasing number of connected devices, an increasing number of sensors in these and an overall need to collect new types of data. In our experience, a common challenge is how to make effective use of data to support development and improvement of software functionality. In this area, existing frameworks are few and for most software product managers the opportunities data provides are also associated with several challenges.

Response: In use case 2 in company A, we studied how machine learning (ML) is used to improve paging. The paging feature is an existing feature in the audio stream that detects when the connection is poor. However, due to the increasing complexity associated with large telecom networks, and competing factors such as latency, resource consumption and number of paging requests, the intention was to explore to what extent the paging feature could be improved by using ML. From a SPM perspective, the use case illustrates the opportunity to have AI technologies complement and even replace human efforts during software development. Also, it shows how ML models can help realize system functionality and perform classification and prediction activities that would be challenging for humans to accomplish.

4.5 Managing Problem Domain Evolution

Challenge: The case companies operate in safety-critical environments where system quality and performance is key. Significant effort goes into continuous monitoring of system to ensure and improve their performance. While it could be argued that quality is important for any system, the systems we studied operate in contexts where failure could lead to severe accidents and even deaths. Therefore, ways in which quality can be assured and continuously improved are critical. At the same time however, internal resources are limited and all companies face challenges with regards to how to increase quality while maintaining, or ideally decreasing, costs involved in this.

Response: In case company C, we studied a use case where the company uses deep learning (DL) models to detect defects in packaging at each client site during processing. The architecture of this use case was presented in [14] where we show how a global model in the cloud is trained with the knowledge gained from local model training at each client site. The learnings from the cloud are fed back and shared to the client sites for inference using transfer learning. The data set consists of packages with different patterns, types and colours and with the DL approach the case company could optimize performance and minimize risks involved in the production line. From a SPM perspective, this allows for an effective way to enhance quality assurance of products while at the same time reduce efforts and costs involved.

4.6 Let the System Figure It Out

Challenge: With the rapidly growing interest in AI, the case companies we studied are looking for approaches that help them use these technologies to explore,

learn and adjust to changes in the environment in which their systems operate. This is of particular interest in contexts characterized by low uncertainty. As a common method, federated learning helps enable large-scale training of models on the device where the data is generated, but with the sensitive data remaining within the data's owner. The approach is generally applicable when the data is evenly distributed across devices. However, in the case companies involved in our study, data is typically not uniform. Also, the data subjects may have different characteristics from one another. The quality of the trained model may then be problematic.

Response: In company D, we studied a use case in which a team used reinforcement learning to explore the reward of introducing a new feature into existing autonomous trucks. In particular, the use case is concerned with monocular depth estimation and in a recent paper we present the software engineering aspects of this case by detailing the ML algorithm, the data sets and the simulations that were used [37]. From a SPM perspective, the reinforcement learning approach allowed for effective exploration of an action space to determine if there was sufficient reward to be accomplished by introducing the monocular depth estimation feature to existing autonomous trucks.

5 SPM4AI: Strategic Digital Product Management in the Age of AI

Software product management is concerned with determining what to build. The goal of this decision process is to maximize the return on the investment of the R&D resources. To accomplish this, the product manager is required to predict what the impact of a function or features on the customer, market and other stakeholders will be. However, predicting the impact of new functionality is far from trivial and traditionally the software product manager simply had to prioritize the content of a release based on their best understanding and assessment. With the emergence of DevOps, we get a new mechanism available: as the release frequency is so high, we can afford to experiment with new functionality before completing it. In this way, DevOps allows for building a slice of new functionality, get it out to some of the customers and use experiments to incrementally add and improve a feature. Experimentation is particularly important in cases where the certainty that a feature will add value is low. Research shows that potentially more than half of all features in a system are never used or used so seldom that the R&D investment was wasted [24]. Experimentation is a powerful approach to address this challenge as we can answer the question on whether functionality adds value with a much lower investment. A second dimension of decision-making is how to realize functionality. Traditionally, all functionality was realized using algorithmic code developed by software engineers. With the emergence of AI, it becomes increasingly feasible to train ML/DL models with available data. These models can then perform classification, prediction as well as other forms of inference.

The challenge of uncertainty and change over time also exists for ML/DL models. In some cases, the input data and the domain in which the system operates is rather static and it is sufficient to train a model once and deploy it. In many situations, however, the context in which the system operates evolves over time. In the context of ML/DL models, there are two basic approaches to accomplish evolution of models. First, one can monitor the performance of an otherwise static ML/DL model. When the performance of the model starts to decrease, this can be used to trigger retraining of the model. This is an effective approach to evolve ML/DL models in changing contexts with an element of human supervision. Although the trigger for retraining may be automated, in most cases there is a human who decides whether a new model goes live or not.

An alternative to retraining models is to use reinforcement learning. In this case, the algorithm is given a state space and an action space. Based on the action the reinforcement learning algorithm takes it receives a reward. Based on this, the algorithm learns, over time, what action is preferred in each situation. In an evolving system, the algorithm continuously spends a small amount of its time exploring. Consequently, when an alternative action is becomes more suitable over time, meaning the reward goes up, the algorithm will learn this and adjust its behaviour.

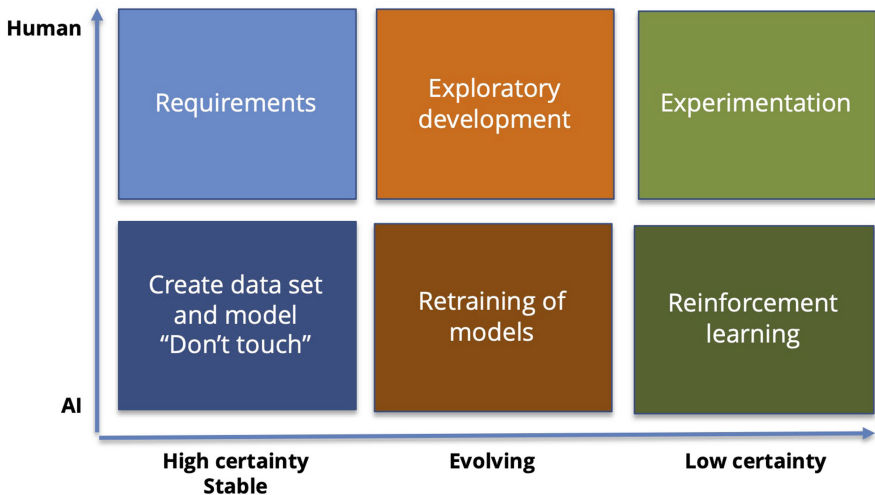


Fig. 1. SPM4AI Framework: six approaches

In Fig. 1, the insights that we developed during our study are summarized. When the functionality prioritized by the software product manager is considered to be stable and we have a high degree of certainty, we can either ask the R&D team to build the functionality based on the requirements or train a ML/DL model if there is data available.

In cases when the context in which the system operates evolves, the system has to respond to these changes. When the functionality is developed by humans, software product managers need to provide updated requirements for the development teams. The challenge is that even if it is obvious that the system needs to respond to changes, it may not be obvious how it should do so. To address this, we propose exploratory development where teams try alternative solutions to figure out the most rewarding path forward.

If the functionality is realized by ML/DL models, the typical approach is to retrain the model using the most recent data. There are challenges around when to retrain, define trigger points and how to ensure that appropriate monitoring is in place. Still, the opportunity to use ML models for managing system evolution is critical for SPM practices going forward as it comes with benefits that are hard to accomplish in traditional software development. If the degree of uncertainty is high to the point that it is not even clear that the functionality should be part of the system, companies need experimental approaches. As we shared earlier in the paper, many features in contemporary systems are never or hardly ever used. The goal of experimentation is to determine whether a new feature should be part of the system at all. If the software product manager decides that a new feature or function should be realized through algorithmic code developed by a team, the suitable approach is to ask the team to conduct A/B experiments. The goal of the A/B experiments is to determine if there is sufficient value for customers or the company providing the system to its customers. In the case the software product manager decides that using an ML/DL model is the best way to realize the feature, reinforcement learning can be an effective approach to determine if there is sufficient reward to be accomplished.

To summarize this section, the role of product manager is to decide what to build in high degrees of uncertainty and a continuously evolving contexts. The framework we present identifies six approaches of realizing functionality that meets the specific constraints for each of the identified situations. In the end, the product manager needs to decide between these approaches based on his or her best understanding of the situation. In general our guidance is to select ML/DL models over algorithm-based development when feasible and to treat new functionality with more uncertainty than what one might believe. Both these guidelines allow for data driven decision making and reduced development efforts.

6 Threats to Validity

As the foundation for our understanding of the impact of digitalization on software product management practices, we reviewed contemporary research on this topic. Based on this understanding, we conducted multi-case study research in collaboration with companies in the embedded systems domain. As our primary data source, we collected data from workshops with key stakeholders within each of the case companies. To address construct validity [22], we shared our understanding of digital transformation, and the impact this has on SPM with all stakeholders involved in our research. With regards to external validity, we view

our research contributions as related to the “drawing of specific implications” and as a contribution of “rich insights” [34]. However, with the opportunity to study companies covering different industry domains we believe that the findings have the potential to be relevant also in other embedded systems companies with similar characteristics as the companies we studied.

7 Conclusion

The role of software product management is key for building, implementing and managing software products. However, few studies explore how this role is rapidly changing due to digitalization and digital transformation. In this paper, we study how key trends such as DevOps, data and artificial intelligence, and digital ecosystems are fundamentally changing current SPM practices. To support this change, and to provide guidelines for future practices, we identify the key challenges that software-intensive embedded systems companies experience with regards to current SPM practices. Second, we present an empirically derived framework for strategic digital product management (SPM4AI) in which we outline what we believe are key practices for SPM in the age of AI.

References

1. Armstrong, M., Wright, J.: Two-sided markets, competitive bottlenecks and exclusive contracts. *Econ. Theor.* **32**, 353–380 (2007)
2. Bekkers, W., van de Weerd, I., Spruit, M., Brinkkemper, S.: A framework for process improvement in software product management. In: Riel, A., O’Connor, R., Tichkiewitch, S., Messnarz, R. (eds.) *EuroSPI 2010. CCIS*, vol. 99, pp. 1–12. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15666-3_1
3. Benbasat, I., Goldstein, D.K., Mead, M.: The case research strategy in studies of information systems. *MIS quarterly*, pp. 369–386 (1987)
4. Bosch, J.: Speed, data, and ecosystems: the future of software engineering. *IEEE Softw.* **33**(1), 82–88 (2015)
5. Bosch, J., Olsson, H.H.: Digital for real: a multicase study on the digital transformation of companies in the embedded systems domain. *J. Softw. Evol. Process* **33**(5), e2333 (2021)
6. Bosch, J., Olsson, H.H., Crnkovic, I.: It takes three to tango: Requirement, outcome/data, and AI driven development. In: *SiBW*, pp. 177–192 (2018)
7. Ebert, C.: The impacts of software product management. *J. Syst. Softw.* **80**(6), 850–861 (2007)
8. Ebert, C.: Software product management. *IEEE Softw.* **31**(3), 21–24 (2014)
9. Ebert, C., Gallardo, G., Hernantes, J., Serrano, N.: DevOps. *IEEE Softw.* **33**(3), 94–100 (2016)
10. Franch, X., Ruhe, G.: Software release planning. In: *Proceedings of the 38th International Conference on Software Engineering Companion*, pp. 894–895 (2016)
11. Fricker, S.A.: Software product management. *Software for People: Fundamentals, trends and best practices*, pp. 53–81 (2012)
12. Gorchels, L.: *The Product Manager’s Handbook 4/E*. McGraw Hill Professional (2011)

13. Gorschek, T., Gomes, A., Petterson, A., Torkar, R.: Introduction of a process maturity model for market-driven product management and requirements engineering. *J. Softw. Evol. Process* **24**(1), 83–113 (2012)
14. John, M.M., Olsson, H.H., Bosch, J.: Towards an AI-driven business development framework: a multi-case study. *J. Softw. Evol. Process* **35**(6), e2432 (2023)
15. Kittlaus, H.B., Fricker, S.A.: *Software product management*, vol. 298. SpringerVerlag GmbH Germany, Berlin (2017)
16. Koch, M., Krohmer, D., Naab, M., Rost, D., Trapp, M.: A matter of definition: criteria for digital ecosystems. *Digit. Bus.* **2**(2), 100027 (2022)
17. Krogstie, J.: Modeling of digital ecosystems: challenges and opportunities. In: Camarinha-Matos, L.M., Xu, L., Afsarmanesh, H. (eds.) *PRO-VE 2012. IAICT*, vol. 380, pp. 137–145. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-32775-9_14
18. Leite, L., Rocha, C., Kon, F., Milojicic, D., Meirelles, P.: A survey of DevOps concepts and challenges. *ACM Comput. Surv. (CSUR)* **52**(6), 1–35 (2019)
19. Liu, Y., Bosch, J., Olsson, H.H., Lantz, J.: An architecture for enabling a/b experiments in automotive embedded software. In: 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC), pp. 992–997. IEEE (2021)
20. Liu, Y., Mattos, D.I., Bosch, J., Olsson, H.H., Lantz, J.: Size matters? Or not: A/B testing with limited sample in automotive embedded software. In: 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 300–307. IEEE (2021)
21. Maglyas, A., Nikula, U., Smolander, K.: What do we know about software product management? - a systematic mapping study. In: 2011 Fifth International Workshop on Software Product Management (IWSPM), pp. 26–35 (2011). <https://doi.org/10.1109/IWSPM.2011.6046201>
22. Maxwell, J.A.: *Qualitative Research Design: An Interactive Approach*. Sage Publications, California (2012)
23. Olsson, H., Sandberg, A., Bosch, J., Alahyari, H.: Scale and responsiveness in large-scale software development. *IEEE Softw.* **31**(5), 87–93 (2013)
24. Olsson, H.H., Bosch, J.: The HYPEX model: from opinions to data-driven software development. In: Bosch, J. (ed.) *Continuous Software Engineering*, pp. 155–164. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11283-1_13
25. Olsson, H.H., Bosch, J.: Going digital: disruption and transformation in software-intensive embedded systems ecosystems. *J. Softw. Evol. Process* **32**(6), e2249 (2020)
26. Olsson, H.H., Bosch, J.: Living in a pink cloud or fighting a whack-a-mole? on the creation of recurring revenue streams in the embedded systems domain. In: 2022 48th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 161–168. IEEE (2022)
27. Parikh, N.A.: Empowering business transformation: the positive impact and ethical considerations of generative AI in software product management—a systematic literature review. *arXiv preprint arXiv:2306.04605* (2023)
28. Pattyn, F.: Enhancing startup success rates: towards a pragmatic framework for product managers (PFPM). In: IEEE 31st International Requirements Engineering Conference (RE) (2023)
29. Putta, A., Paasivaara, M., Lassenius, C.: Benefits and challenges of adopting the scaled agile framework (SAFe): preliminary results from a multivocal literature review. In: Kuhrmann, M., et al. (eds.) *PROFES 2018. LNCS*, vol. 11271, pp. 334–351. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03673-7_24

30. Rautiainen, K., Lassenius, C., Sulonen, R.: 4CC: a framework for managing software product development. *Eng. Manag. J.* **14**(2), 27–32 (2002)
31. Rochet, J.C., Tirole, J.: Two-sided markets: an overview. Institut d'Economie Industrielle Working Paper, pp. 1–44 (2004)
32. Stallinger, F., Neumann, R.: From software to software system products: an add-on process reference model for enhancing ISO/IEC 12207 with product management and system-level reuse. In: 2012 38th Euromicro Conference on Software Engineering and Advanced Applications, pp. 307–314 (2012). <https://doi.org/10.1109/SEAA.2012.41>
33. Svahnberg, M., Gorschek, T., Feldt, R., Torkar, R., Saleem, S.B., Shafique, M.U.: A systematic review on strategic release planning models. *Inf. Softw. Technol.* **52**(3), 237–248 (2010)
34. Walsham, G.: Interpretive case studies in is research: nature and method. *Eur. J. Inf. Syst.* **4**(2), 74–81 (1995)
35. van de Weerd, I., Brinkkemper, S., Nieuwenhuis, R., Versendaal, J., Bijlsma, L.: Towards a reference framework for software product management. In: 14th IEEE International Requirements Engineering Conference (RE 2006), pp. 319–322 (2006). <https://doi.org/10.1109/RE.2006.66>
36. Wohlin, C., Aurum, A.: Criteria for selecting software requirements to create product value: an industrial empirical study. In: Biffl, S., Aurum, A., Boehm, B., Erdogmus, H., Grünbacher, P. (eds.) *Value-Based Software Engineering*, pp. 179–200. Springer, Heidelberg (2006). https://doi.org/10.1007/3-540-29263-2_9
37. Zhang, H., Li, J., Qi, Z., Aronsson, A., Bosch, J., Olsson, H.H.: Deep reinforcement learning for multiple agents in a decentralized architecture: a case study in the telecommunication domain. In: 2023 IEEE 20th International Conference on Software Architecture Companion (ICSA-C), pp. 183–186. IEEE (2023)


Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Experimentation in Early-Stage Video Game Startups: Practices and Challenges

Henry Edison¹(✉) , Jorge Melegati² , and Elizabeth Bjarnason³ 

¹ Blekinge Institute of Technology, Karlskrona, Sweden
henry.edison@bth.se

² Free University of Bozen-Bolzano, Bolzano, Italy

³ Lund University, Lund, Sweden

Abstract. Experimentation has been considered critical for successful software product and business development, including in video game startups. Video game startups need “wow” qualities that distinguish them from the competition. Thus, they need to continuously experiment to find these qualities before running out of time and resources. In this study, we aimed to explore how these companies perform experimentation. We interviewed four co-founders of video game startups. Our findings identify six practices, or scenarios, through which video game startups conduct experiments and challenges associated with these. The initial results could inform these startups about the possibilities and challenges and guide future research.

Keywords: experimentation · video game startups · challenges · gaming startups

1 Introduction

Over the last 40 years, video games have increasingly replaced traditional games as leisure activities and have disrupted how we spend our leisure time. The video game market has become an established and ever-growing global industry for over two decades. In 2022, the global video market was worth USD 42.9 billion, and the revenue is expected to grow with an annual growth rate of 8.74%¹. Originally, video games refer to the games that do not require a microprocessor and use analogue intensity signals displayed on a cathode ray tube (CRT) [17]. The availability of new imaging technologies, such as consoles, home computers, Virtual Reality (VR) devices, etc., has made the idea of video games more conceptual and less tied to a specific technology [5].

Developing a successful video game is a very demanding and complex process. It involves expertise from various disciplines, e.g. software/game development, arts, animation, sound engineering, etc., which may increase the complexity of

¹ <https://www.statista.com/statistics/292516/pc-online-game-market-value-worldwide/>.

communication and coordination [10]. Furthermore, it is unclear whether a game will succeed in the market, which poses a major risk to game publishers when investing in new game development projects. Unlike other software startups, video game startups do not build technological solutions to solve real problems. Instead, they combine art, science, and craft to offer fun, entertainment, and experience through the games [2, 11]. Yet, these requirements have no metrics to be applied, yet they must be validated at each stage of the development process.

An effective adoption and implementation of experimentation is a staged process [13]. In this study, we aim to gain insights into how video game startups approach experimentation to develop games. To guide the study, we explore the research question: *How do video game startups use experimentation in practice?*

2 Background and Related Work

In innovative endeavours, the required knowledge for success is generally unknown [9]. Thus, experimentation is particularly useful for acquiring knowledge and reducing uncertainty. Experimentation is an approach based on continuously identifying critical assumptions, transforming them as hypotheses, and prioritising and testing them with experiments to support or refute them [12]. However, most startups persist with the original ideas rather than experimenting [6, 7, 14].

While research in game startups exists, they are limited to mobile game development. For example, Vanhala et al. [16] analysed six Finnish mobile game startups and found that human capital is the most important element in their business models. Moreover, the key challenge is to raise the awareness of game players. Kasurinen et al. [8] showed that game developers are generally pleased by the tools available to experiment with the concept and build prototypes.

Research also shows that the iterative and incremental nature of agile methods positively impacts communication, game quality, and the ability to find the fun aspects of the mobile game features [10]. In contrast, the agile principle of embracing changes increases the pressure to meet the deadline [1]. Mobile game startups should be cautious in considering the minimum viable product concept. The first version of a game artefact released to the market needs to be of sufficient quality to attract and lock in users for an adequate amount of time to allow for further development of the game [15]. This study aims to complement existing research by investigating how video game startups conduct experimentation.

3 Research Methodology

We performed semi-structured interviews [3] to gain insights into how video game startups conduct experimentation. Interview candidates were identified by the first author collaborating with Blekinge Business Incubator (BBI) in Karlskrona, Sweden. The first interview was with a business coach in the incubator, who provided a list of founders of independent (indie) and internal video game startups

operating inside larger companies. The interviews were held and recorded in a video conferencing system (Microsoft Teams), each lasting between 60 and 90 min. The profiles of the interviewees are shown in Table 1. The audio recordings were transcribed and analysed using thematic analysis [4]. The transcripts were sent back to the interviewees for follow-up questions and clarification.

Table 1. Overview of interviewees

ID	Company name	Company age (years)	Number of people	Type of company	Game genre	Experience (years)	Role
A	BBI		–	Incubator	–	10	Business coach
B	Mana Brigade	3	5	Indie startup	Adventure	3	Founder/CEO
C	The Station	10	35–40	Internal startup	Simulation	14	Founder/Game director
D	Blackdrop Interactive	8	4–5	Internal startup	Warfare	8	Founder

4 Results

This section reports our findings by describing six experimentation scenarios. All quotes and information herein are derived from the interview transcripts.

4.1 Technical or Digital Prototyping

Our interviews reveal that, in the early stages, the main challenge of game development lies not in the ideation process but in the execution and making the game work. Hence, the first purpose of experimentation is to assess the technical feasibility of the team to develop the game. The game’s initial idea is usually outlined in a game design document and describes the game at a high level from the user’s perspective. The team builds prototypes using a 3D engine, e.g., Unity, to test the game’s complexity and scope. In Mana Brigade, a slightly different approach was taken. This company started out performing experiments with a marching cube algorithm². This algorithm was then implemented in Unity, and the user experience was tested using VR devices.

All interviewees agree that technical experimentation is crucial to evaluate their capability to build the game. For example, if they can solve all problems to build a game or need key people with certain skills and expertise. Technical experimentation also showcases their capabilities to potential investors or publishers.

² Marching cubes is an algorithm to extract a 2D surface mesh from a 3D volume.

4.2 Controlled Game Tests

Game startups also experiment with external stakeholders, such as end users or players, to evaluate whether they understand the game's concepts and mechanics. In the case of The Station, they hired external game companies to test their game: “[The external video game companies] bring in players. We have a questionnaire that we want them to answer that they rate the game [like] ‘Was there anything unclear? What did you not like? What did you like?’” (Interviewee C)

4.3 Mock Reviews

In the case of The Station, they asked game journalists to write a mock review and to give a score of their game compared to other games in the same genre. The score was used as an early indicator of what could happen when the game was released. In the case of Mana Brigade, they mentioned that it does not use this approach due to a lack of funding.

4.4 Presenting and Pitching in Game Conferences

Presenting and pitching new games in video game conferences is a good opportunity to validate assumptions about the game, e.g., the basic idea and its potential market. In these events, video game startups can meet and talk to publishers, investors, or game scouts to get investments from them to build the game. Mana Brigade's first experimentation with external stakeholders was competing in a game competition in 2021. “For the first iteration, we want it to be multiplayer, and [we want] to explore dungeons. It's like awesome, like real-time events. [But] we got feedback from the [judges] ‘This doesn't make sense.’ So we took that year to iterate on it, and then we wanted to do like it was still single player, but it was still crafting and then adventuring.” (Interviewee B)

However, explaining and convincing the game concepts and design to publishers is a big problem. Video game startups need to find ways to explain their game and, at the same time, to find the right publishers: “[Publishers] get bombarded with hundreds of game ideas they must go through to find that one good game... One publisher wants a game design document, not a PowerPoint. They don't care about the pictures, [while others] want many. It's very hard to know what they want.” (Interviewee B)

4.5 Social Media Engagement

The interviewees expressed that they could use social media platforms, i.e., YouTube or Instagram, to experiment and gain user feedback. For example, by releasing screenshots, images, videos or tutorials on social media and measuring gamers' reactions to these. However, this may not work for indie game startups. They must balance the effort and resources between developing the game and actively maintaining communication with the community and the users.

4.6 Early Release of Vertical Slice

Releasing a vertical slice³ on video game platforms like Steam for user testing may allow game startups to build a player base. It may also give them some small revenue to improve the game, but it could harm their reputation. Besides that, they need to find the right audience for their games: “*The game industry is so big... maybe 100 [new games are published] every day on Steam. It’s hard to reach and find your audience and see your game. There is so much information [on Steam], and many games [can easily] get drowned.*” (Interviewee A)

5 Discussions and Conclusions

Table 2 summarises the six practices we identified and their associated challenges. Some of the practices are present in other contexts, e.g. prototyping. Some are adapted to the context of games, e.g. controlled game tests and early release, while some are specific to the game industry, e.g. mock reviews by journalists and presentations in game conferences.

Table 2. Experiment practices and challenges in video game startups

Practice	Purpose	Challenges
Technical/Digital Prototyping	Understanding the game complexity and team capability	Missing skill-sets and expertise in-house
Controlled game test	Understanding if users understand the game concepts and mechanics	Funding to hire professional game testers
Early (vertical) release	Build user base and get early revenue	Find the right audience, maintain the reputation
Social media engagement	Build user base	Need high effort
Mock reviews by game journalists	Estimate the review score	Funding to hire professional game journalist
Presenting and pitching the game	Understanding the market potential and securing funding	Explaining the game’s concepts and design to publishers

The identified challenges can be related to the experimentation inhibitors experimentation identified by Melegati et al. [13]. Missing skill sets and expertise and lack of funding to hire game testers or journalists relate to the scarcity of

³ A vertical slice is a fully playable portion of a game that shows its developer’s intended player experience.

technical and development resources. The need for early releases is associated with time pressure and over-focus on customer base growth in the early phase. However, the difficulty of explaining a game's concepts to publishers might be considered a specific challenge of video game startups. It could be classified as an inhibitor to a valid experiment, as described by Melegati et al. In summary, our study describes the particularities of video game startups and provides evidence to support an existing model in the literature.

This study poses a first step to understanding experimentation within gaming startups. Next, additional video game startups will be studied to further expand on their experimentation practices. We will also expand beyond studying startups that develop games for specific platforms, such as consoles and VR, including other platforms, such as smartphones and tablets. By contrasting and comparing the results, we can improve the generalisability of the findings. Future research could also investigate gaming startups' use of novel technologies, such as artificial intelligence and how these affect their experimentation.

Acknowledgement. This work has been supported by ELLIIT; the Swedish Strategic Research Area in IT and Mobile Communications.

References

1. Borg, M., Garousi, V., Mahmoud, A., Olsson, T., Stålberg, O.: Video game development in a rush: a survey of the global game jam participants. *IEEE Trans. Games* **12**(3), 246–259 (2020)
2. Crawford, C.: *The Art of Computer Game Design*. McGraw-Hill (1984)
3. Creswell, J.W.: *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publication (2009)
4. Cruzes, D.S., Dyba, T.: Recommended steps for thematic synthesis in software engineering. In: 2011 International Symposium on Empirical Software Engineering and Measurement, pp. 275–284. No. 7491. IEEE (2011)
5. Egenfeldt-Nielsen, S., Smith, J.H., Tosca, S.P.: *Understanding Video Games, the Essential Introduction*. Routledge (2020)
6. Giardino, C., Wang, X., Abrahamsson, P.: Why early-stage software startups fail: a behavioral framework. In: *Proceedings of ICSOB*, vol. 182, pp. 27–41 (2014)
7. Gutbrod, Matthias, Münch, Jürgen., Tichy, Matthias: How do software startups approach experimentation? Empirical results from a qualitative interview study. In: Felderer, Michael, Méndez Fernández, Daniel, Turhan, Burak, Kalinowski, Marcos, Sarro, Federica, Winkler, Dietmar (eds.) *PROFES 2017. LNCS*, vol. 10611, pp. 297–304. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-69926-4_21
8. Kasurinen, J.: What do game developers expect from development and design tools. In: *EASE 2013 Proceedings* (2013)
9. Kerr, W.R., Nanda, R., Rhodes-Kropf, M.: Entrepreneurship as experimentation. *J. Econ. Perspect.* **28**(3), 25–48 (2014)
10. Koutonen, J., Leppänen, M.: How are agile methods and practices deployed in video game development? A survey into finnish game studios. In: *Agile Processes in Software Engineering and Extreme Programming*, pp. 135–149 (2013)
11. Lewis, C., Whitehead, J.: How are agile methods and practices deployed in video game development? A survey into finnish game studios. In: *Proceedings of Workshop on Games and Software Engineering*, pp. 1–4 (2011)

12. Lindgren, E., Münch, J.: Raising the odds of success: the current state of experimentation in product development. *Inf. Softw. Technol.* **77**, 80–91 (2016)
13. Melegati, J., Edison, H., Wang, X.: XPro: a model to explain the limited adoption and implementation of experimentation in software startups. *IEEE Trans. Software Eng.* **48**(6), 1929–1946 (2020)
14. Pantiuchina, J., Mondini, M., Khanna, D., Wang, X., Abrahamsson, P.: Are Software Startups Applying Agile Practices? The State of the Practice from a Large Survey, vol. 283, pp. 167–183 (2017)
15. Roshan, M., Hekkala, R., Tuunainen, V.K.: Dynamics of success for mobile games: the case of early-stage startups. In: *ECIS 2019 Proceedings* (2019)
16. Vanhala, E., Kasurinen, J.: The role of business model and its elements in computer game start-ups. In: *Proceedings ICSOB*, pp. 72–87, Cham (2014)
17. Wolf, M.J.P.: *The Game Video Explosion*. Greenwood Press (2008)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Software and Business Co-Development



DevOps Challenges and Risk Mitigation Strategies by DevOps Professionals Teams

Nasreen Azad^(✉) 

Department of Software Engineering, Lappeenranta, Finland
nasren.azad@lut.fi

Abstract. DevOps is a team culture and organizational practice that eliminates inefficiencies and bottlenecks in the DevOps infrastructure. While many companies are adopting DevOps practices, it can still be risky. We conducted 26 interviews with DevOps professionals around the globe and found four major risks associated with DevOps practices: Organizational risks (Intra-organizational collaboration and communication, strategic planning), Social and cultural risks (Team Dynamics, Cultural shift), Technical risks (Integration, Build and test automation), Ethics and security breaches in DevOps environment (Ethical risks, Data collection ethics, Ethical decision making). Our research also identified several risk mitigation strategies namely continuous testing, using infrastructure as code, security audit and monitoring, disaster recovery planning, cross-functional training, proper documentation, continuous learning, continuous improvement etc. that companies can adopt for better performance and efficiency.

Keywords: DevOps · DevOps practice · DevOps risks · DevOps risk mitigation strategies · Qualitative research

1 Introduction

In traditional software development, separate teams handle operations, security, and quality assurance. However, conflicts between development and operations teams can arise while delivering software [5]. Upon observing the software development process, it becomes clear that operations require a high level of security and stability, while simultaneously expecting developers to minimize changes to upcoming products. Nevertheless, developers must frequently work on new features, upgrade existing ones, and make changes to meet customers' evolving needs with confidence [5]. As development teams strive to release new versions faster, operations teams may be reluctant to accept many changes in old versions, leading to conflicting situations [2]. These sort of conflicts can reduce the software development process and makes the release slower [5].

DevOps is an emerging concept and is a blend word of operations and development that is used to eliminate the gaps between Dev and Ops teams so that collaboration and communication can flow clearly with the sharing approaches

for the software development life cycle [2]. According to Debois, DevOps concepts work for medium to large-size organizations and help companies to bridge the gaps between teams [8]. DevOps is a mixture that improves collaboration and communications to solve critical problems for the software development phase [1]. While working on software development, the teams could meet many challenges and risks and DevOps provides support to eliminate the conflicting issues between teams [1]

Implementing continuous deployment of software has opened up new opportunities for companies, but it has also presented numerous challenges and risks [17]. When a company decides to adopt DevOps, they may encounter various challenges in different stages of the software cycle, such as organizational, cultural, social, technical, and managerial challenges [2]. Since the adoption of DevOps can be a difficult process for companies, they can support the process by incorporating technological changes, implementing new processes, hiring trained personnel and consultants, and being open to innovation. The adoption of DevOps in a company is a distinct process that produces many risks and mitigation strategy impact multiple aspects of DevOps practices [2].

However, there are limitations of the DevOps literature as there are a small number of research studies dedicated to DevOps risks and mitigation strategies for the software development cycle. Moreover, there are no clear risk mitigation strategies described in the literature. Therefore, we are interested in focusing on understanding the various risk factors along with the mitigation strategies proposed by the industry professionals in using DevOps in IT organizations. The author believes that identified risks and risk mitigation strategies will be a great contribution to companies, and DevOps practitioners to understand how to perform effective risk management in a DevOps environment.

The remaining of this study is organized as follows. Section 2 presents DevOps concepts, DevOps implementation and benefits, DevOps risks and risk mitigation process, and their related literature. It is followed by the description of the empirical data collection and the research process in Sect. 3. Section 4 presents the results, Sect. 5 discusses their impacts, and Sect. 6 concludes the study.

2 Related Work

2.1 DevOps Concept

Professionals describe DevOps as a software engineering culture, work practice or even a philosophy. If we observe the scientific community, different views, perceptions and stances have been developed and suggested regarding DevOps. DevOps describes how cross-functional teams work together to build, test and release faster software more reliably [18]. Automation plays a vital role in DevOps operations as its goal is to improve collaboration between two teams in terms of software development.

2.2 DevOps Implementation and Benefit

Organizations are increasingly adopting DevOps practices to enhance their software delivery process [23]. By effectively implementing and adopting DevOps principles, the gap between development and operations teams can be minimized. The development process triggers software deployment, which is crucial for software organizations to move software into production [8]. The key aspect of DevOps in an organization is to ensure continuous delivery and deployment, resulting in faster software delivery cycles [10]. As a result, DevOps has become an essential part of modern software development, providing organizations with a competitive edge and enabling them to stay ahead in the market.

Krey et al. [15] have identified six major challenges faced by small and medium-sized enterprises in DevOps implementation: costs, risks, scope, quality, business value, and time. However, a lack of communication among teams can be a major contributor to unsuccessful DevOps adoption. Operations teams have specific responsibilities, and they often don't pass or monitor different performance metrics that could help developers execute tasks [21].

Companies are increasingly adopting DevOps practices in response to customer and user expectations for software applications that meet their needs [13]. To meet this demand, organizations are striving to release frequently and deploy faster, but this requires an efficient process environment and proper utilization of resources. DevOps helps address miscommunications and gaps in the process with four guiding principles: automation, culture, collaboration, and measurement [13]. Gupta et al. [13] also identified four variables that impact the implementation process: source control, automation, cohesive teams, and continuous delivery. By addressing these factors, organizations can successfully implement and adopt DevOps practices.

2.3 DevOps Risks and Risk Mitigation

Effective collaboration between development teams and operations teams is crucial for successful software development and deployment. To facilitate this, it is important to have a common set of tools used by both teams, as using different toolsets can create problems and inefficiencies in the collaboration process [6]. Communication between the Dev and Ops teams is also of utmost importance, as lack of communication can lead to delays in the operating process of both teams [21]. DevOps leverages a variety of tools to streamline the software development process. However, the COVID-19 pandemic has forced most of the work to go remote, which has had a significant impact on the working process [20]. It is important to note that electronic tools alone cannot solve all problems and some issues are best addressed in person. Furthermore, integrating different tools can be challenging and require additional maintenance and execution efforts [5].

Companies can employ various strategies to effectively address risks and challenges. One such strategy is to move away from the traditional Dev and Ops mindset and embrace continuous delivery practices. Adopting microservices-based infrastructure and architecture, implementing test automation techniques,

prioritizing tools, delegating release ownership to teams, and fostering a culture of continuous learning are also effective strategies. Jones et al. [14] recommend the introduction of job crafting as a means to help DevOps professionals achieve their personal goals. Job crafting is an individualized design process that allows employees to proactively modify job characteristics to align personal growth with work objectives. Through job crafting, employees gain greater control over their tasks, determine how their work is perceived, and shape the social context and relationships within the workplace [4]. According to Jones et al. [14], task, relational, and cognitive job crafting can significantly enhance work performance while adopting DevOps in companies. Liete et al. [16] suggest three approaches for implementing DevOps adoption in companies: department collaboration, DevOps teams, and cross-functional teams.

2.4 Research Questions

The aim of this paper is to identify the challenges and risks that IT companies face when adopting DevOps, and how they mitigate these risks by implementing various strategies. We have conducted in-depth interviews (N=26) with DevOps professionals from different companies around the world to investigate these issues. As a result, we will try to answer the following research questions in this paper:

RQ1: What are the risks associated with DevOps practices in organizations?

RQ2: What strategies are used by professionals for risk mitigation?

3 Research Approach

3.1 Data Collection

Throughout our research, we had the privilege of interviewing multiple accomplished DevOps professionals in order to gather valuable data. Our research methodology involved conducting thorough interviews to pinpoint prevalent obstacles and potential hazards that professionals face, examine professional practices, address security concerns, and deeply explore the ethical considerations within DevOps teams. To ensure our interviews were comprehensive, we created a set of 18 questions organized into three themes: challenges and risks overview with mitigation, security risk and mitigation, and team ethics and mitigation strategies from technical, social, and cultural viewpoints.

During the course of the study, respondents represented companies ranging from 80 to 15,000 employees. The respondents held various positions within their respective organizations, including Head of Technology, Tech Lead, Scrum Master, Site Reliability Engineer, DevOps Engineer, Software Specialist, Business Analyst, Cloud Engineer, Technical Project Manager, and Software Engineer. With working experience in the software development industry ranging from one to twenty years, respondents were contacted via email for participation in the interview. The interviews were scheduled for a duration of thirty minutes,

during which in-depth questions were asked, focusing on specific areas of DevOps practices. The researcher worked diligently to ensure that the data collected was accurate and relevant to the study's objectives.

During the interview process, we ensured that each interviewee provided their consent to being recorded. For those who declined to be recorded, we respectfully opted to take notes instead. In total, we conducted 26 interviews with distinguished DevOps professionals occupying diverse roles across numerous companies. These interviews were conducted during the first quarter of 2023, specifically from March to April. Subsequently, we performed a comprehensive analysis of the findings based on the interviews. The 26 interviewed individuals represented 26 distinct companies, which we labeled with different alphabets in the presentation of our results.

3.2 Data Analysis

For analyzing the data, we have used the Gioia method presented by Gioia et al. [12]. An iterative process has been followed which ensures the repetition of steps for the data analysis. We have followed open coding for extracting data from the interviews. As a guideline, we have followed Strauss and Cobin [22] for assigning codes for the analysis. We started the coding process with the interview transcripts, then we marked specific areas and assigned the codes suggested by [19]. For the first-round coding, we used the research questions as guidelines. From the empirical data, we checked what are the similar codes in the various segments of the data. Then we checked the dissimilarities present in the codes and identified those codes.

In our research, We utilized constant comparison and followed the grounded theory approach [22]. We have prepared a table that showcases the coding activities created from the interview data, providing an explicit understanding of the coding process. The table includes a detailed list of codes, their corresponding descriptions, and quotes from professionals. An exemplary table called Table 1 illustrates the coding activities. This table can be used as a reference to gain insights about the coding methodology.

After the first coding ended, we moved to the second phase of coding. In the second phase, we have started categorizing the first phase codes. According to Charmaz, to create second-order codes for concepts it is necessary to categorize the first-phase codes [7]. Then we merged the first-phase codes with second-phase codes [11]. To make the data analysis process accurate we have also used memoing techniques. Memoing helped us to understand more insights and perspectives of professionals' views regarding critical success factors and organizational practice. A total of 910 pages were generated from the interview data transcription. According to our understanding, we have used an iterative process for data analysis [22].

In the third phase of the data analysis, we have aggregated the themes into four main aggregate categories including Organizational risk, Cultural and Social risk, and Technical risk and Ethics and security breach risk. In Fig. 1, we have shown the data analysis process with themes.

Table 1. Coding used for interview data

List of codes	Description of codes	Quotes from Professionals
Lack of tacit knowledge	The knowledge base for the system is not strong (tacit knowledge), a knowledgeable person left might impact the company negatively. Losing one key person may ruin the whole process	“In our team different skillset people are working. When someone goes they also take the expertise and knowledge with them, which creates difficulties in teamwork.”
Miscommunication between clients and developers	communication between the clients and developer makes the project run smoothly or otherwise creates miscommunication and unclear perception	“Miscommunication is a drawback for DevOps practices when there is a communication gap that leads to a project failure and makes the project risky to execute.”
Security in the DevOps environment	DevOps security is a set of practices, tools, and cultural approaches that bring together software development, software operations, and security all together to make the process faster and more secure	“To make a project successful it is important to maintain the security from the very beginning of the development process”
Human error on DevOps risks	Human errors are one of the most unpredictable situations for any DevOps team which might create several risks for DevOps environment.	“Human error is difficult to eliminate but if teams maintain some steps then there will be less human error ”
Handling ethical issues while working in teams	DevOps team members need to have the appropriate knowledge and training to understand and address ethical issues that may arise in operations	“It is essential to have proper training and knowledge while working in DevOps teams. The companies have training for team members so that they know how to handle difficult situations”

4 Results

In this section, we will highlight different risk factors associated with technical, organizational, social, and cultural risks while practicing DevOps in teams and organizations. We will also discuss how the professionals handle several DevOps implementation and adoption risks while working in teams and how the risks are mitigated.

4.1 Organizational Risks

Intra-organizational Collaboration and Communication. Recognizing a lack of understanding about the project among team members is essential. Miscommunication caused by unclear project knowledge among those outside of IT teams or the project can significantly jeopardize its success. Additionally, our research indicates that poor communication between clients and developers poses another risk to project success. Inadequate communication creates challenges, misunderstandings, and unclear perceptions, making it imperative to prioritize clear and effective communication throughout the project's development.

A professional quoted:

“In our teams, there are sometimes miscommunications, and due to that DevOps practices get hampered (Development and Operations) and lack of collaboration between clients and developer teams make the process risky, improper communication creates difficulties for better outcomes”.

Strategic Planning. Based on the extensive research by Azad and Hyrynsalmi Azad and Hyrynsalmi [2], the product management team is responsible for maintaining the business requirements, while the technology team handles the technical requirements, emphasizing the need for meticulous planning related to resources, initiatives, and budget for the overall software process. It is crucial for the IT and business plan to share similar goals and objectives. Adopting continuous development and continuous delivery would ensure top-notch quality of the product. Therefore, strategic planning should prioritize company pressure, change management, meeting deadlines, and reducing the time to market Azad and Hyrynsalmi [2].

Our findings suggest that improper allocation of budget for the toolset is a risk for DevOps practices. The budget allocation for toolsets is important because wrong choices create risks for the project. According to professionals risky change and development are challenging for the teams. People in the team are reluctant to new changes as those are uncomfortable and people fear changes.

A professional quoted that:

“Risk mitigation through automated testing and quality assurance is essential for the development process. If automated tests are in place, a developer can immediately get feedback about their newly written programs/features. Then the process becomes less risky”.

Quality assurance acts as a bridge between development and operations teams and supports developers by testing new iterations in real-time with continuous quality checks to keep the testing cycle running smoothly.

Another professional quoted that:

“Balancing security and risk management for the DevOps process is crucial. For good balancing the team needs to make sure that they do not release anything if not properly tested”.

4.2 Social and Cultural Risks

Social and cultural risk factors are one of the leading factors for DevOps risks in the organization [2]. Below we discuss the team dynamics and social and cultural shifts risks.

Team Dynamics. In teams when there is a lack of tacit knowledge then the knowledge base is not strong. If a knowledgeable person leaves, it might impact the company negatively specifically the team dynamics might be hampered. Losing one key person may ruin the whole process and create a setback in the working environment.

A professional quoted that:

“When a team has skilled and knowledgeable people with a diversified culture that helps the team to progress better. A sudden change like someone leaving the team might slow down the process as DevOps teams are connected with each other and that’s the way the team progress”.

Cultural Shift. When the team is reluctant to accept organizational culture that impacts DevOps practices hugely. According to the professionals, security must be considered a part of DevOps from the beginning. The team should make a list of DevOps best practices document and follow strictly and avoid; discussing sensitive information in public places can support a good culture. This makes the process less risky and impacts positively as an organizational culture.

A professional quoted that:

“Lack of collaboration and organizational culture does not help for better building products for clients. The company culture should be collaborative, flexible, and supportive. To make it secured from the beginning DevSecOps should be a part of the process”.

4.3 Technical Risks

There are several technical risks associated with DevOps practices. Some of those include improper code review by team members, security in a DevOps environment, and human error as a DevOps risk.

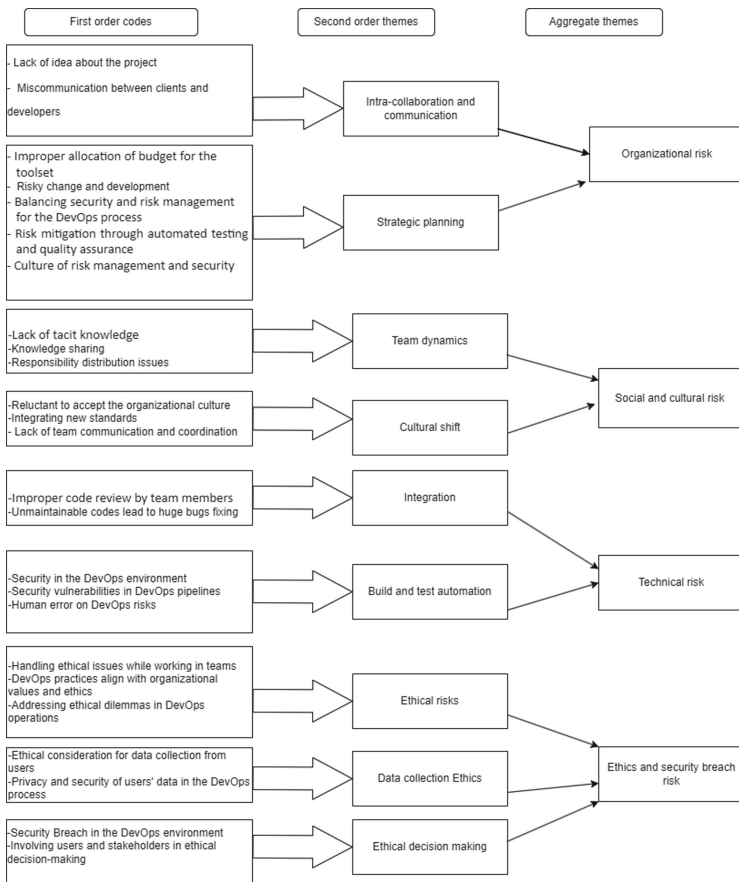


Fig. 1. Themes from data analysis

Integration. Continuous integration is essential for doing several automated actions that help the system work together for the pipeline. Some of the pipeline stages include package generation, automated test execution, code verification, and deployment for the production and development environments. The developers are the responsible actors for defining pipeline structures. On the other hand, operators are responsible for defining collaboration for deployment phases. Developers are also responsible for the continuous integration. When there is an improper code review by team members that impacts the review process hugely. When developers take shortcuts and input unmaintainable codes to fix issues by ignoring the consequences they need to handle a lot more bugs and issues later on.

Build and Test Automation with Security. DevOps security is a set of practices, tools, and cultural approaches that bring together software develop-

ment, software operations, and security all together to make the process faster and more secure. Security in the DevOps environment is one of the vital things to consider for software development. According to the professionals having a proper DevOps architecture and plan, writing test code while developing software, and automated tests should be from the beginning and CI/CD stages - Development, Staging and production.

A professional quoted that:

“Uh, of the project experiences within the company they at first understand the requirements and set up the tools which are actually secured. So the important thing is that selection of the tools that make a secured environment for the development process”.

According to our findings, the professionals stated that security vulnerabilities in DevOps pipelines are risky for the companies. Security vulnerabilities include missing data inscription, missing authentication for critical functions, and buffer overflows with insecure interactions between software companies. Whatever the developer has done and if the test is an improper code review, it is the number one risk for the process.

A professional quoted that:

“For maintaining security vulnerabilities, developers need to check if the web service is running and the Azure function can send requests and get the response back each hour. There should be access restrictions so only certain IPs are allowed if that is required”.

Human errors are one of the most unpredictable situations for any DevOps team which might create several risks for the DevOps environment. There are many steps as a part of DevOps work. People may forget to test certain codes or follow best practices. Maybe one port remains open by mistake, Data Storage is open to public access, Databases does not have IP restrictions, forgets to stop an expensive during holiday/weekend, no cost tracking of the cloud services. These errors could impact the development process hugely.

4.4 Ethics and Security Breach in DevOps Environment

Ethical Risks. Handling ethical issues while working in teams is considered as one of the most important aspects of working in a DevOps environment. DevOps team members need to have the appropriate knowledge and training to understand and address ethical issues that may arise in operations. According to the professionals, DevOps practices align with organizational values and ethics helping the teams to work efficiently.

A professional quoted that:

“DevOps practices align with our organization’s values and ethical principles and require timely release features, Deployment frequencies, Time to recover in case of any issues, data protection, and scalability. ”.

Data Collection Ethics. Data collection ethics is essential for any software development process. Privacy and security of users' data in the DevOps process is important. To maintain privacy and security it should be aligned with the company's values, culture, and security checks.

A professional quoted that:

“For ethical considerations, a company should take into account where, and how to collect, store, and analyze data in our DevOps operations”.

Ethical Decision Making. To maintain the issues with security breaches in the DevOps environment, even before starting a project there should be a secure architecture and make sure the system has been implemented according to the architecture. That makes the system secure. The professionals stated that addressing ethical dilemmas in DevOps operations is something to consider from the beginning of the software development process. This is a matter of team discussion including team members, managers, clients, or maybe other teams as well. Everyone should work as a team and be aligned with the company's business and ethical values.

A professional quoted that:

“Involve users and other stakeholders in ethical decision-making processes related to our DevOps operations is essential. A good communication can solve most of the issues. ”

4.5 Risk Mitigation Strategies by Professionals

To mitigate risks and improve performance, there are various approaches that professionals can adopt. Respondents have highlighted different strategies that can assist in managing organizational risks. According to the research findings some of the risk mitigation strategies could be continuous testing, using infrastructure as a code, security audit and monitoring, disaster recovery planning, cross-functional training, proper documentation, continuous learning, continuous improvement, making process visible to the team members, prioritize personnel so they feel valued, enforce security policy, introduce DecsecOps, involvements of experts from outside, and improved management strategies. In the example Table 2, we have given a short list of risks and risk mitigation strategies proposed by the professionals.

It is imperative to establish a comprehensive framework that can effectively address the issues of security and ethics. To achieve this, it is crucial to facilitate effective communication and establish a robust system of governance. An effective security system or a set of cybersecurity approaches should be implemented to ensure that the security processes are straightforward, transparent, and comprehensible. The security process should encompass a wide range of issues, including code review, access restrictions, and management configuration, among others.

In order to produce a well-secured application, it's crucial for DevOps teams and security teams to work closely together. This collaboration helps to ensure that robust policies and effective tools are implemented to protect the application from potential security threats. By working together, DevOps and security teams can identify potential vulnerabilities in the application and take proactive steps to address them. Additionally, this collaboration can help to streamline the development process by incorporating security measures early on, reducing the likelihood of costly delays or security breaches down the line.

According to the respondents, the challenges for DevOps adoption is insufficient knowledge in industries and the engineers also have a knowledge gap. Though they might have some strong understanding or knowledge or background in some specific part of the software development but DevOps practices still needs to be understood by many of them. DevOps needs proper communication with software developer. The engineers witness that sometimes a developer only working on his coding but when deployment comes, he doesn't have really much idea what's happening in the back end or in the cloud system and also the automation is unclear to him.

5 Discussion

5.1 Key Findings

This study addresses two aspects of DevOps. Firstly, the risk factors identified by industry professionals and, secondly, risk mitigation strategies for DevOps risks. Our findings also discuss security issues and organizational ethics along with risk factors in DevOps operations. From the interviews with professionals, we have learned several DevOps practical risks faced by organizations. However, these risk factors are not universal. These are the professional's own views regarding the risks and the ways to mitigate them when necessary.

According to our findings, there are four major risk factors including organizational risk factors, social and cultural risk factors, technical risk factors, and ethics and security breach factors.

Misunderstanding between Development and Operations teams poses a significant obstacle to the success of the DevOps process. According to research, there is a lack of coordination among team members when working together [1, 2, 15, 21]. This lack of communication can hinder the adoption of DevOps, making the process unsuccessful [1, 15]. One of the most significant risks faced by DevOps teams is the need to balance performance and the speed of releases [3]. Professionals have reported that fast release cycles can enhance performance while reducing the time required for development [3].

Based on the feedback received from participants, it is clear that implementing DevOps in a company can be a difficult and risky task, which may result in an unsuccessful implementation. Employees often struggle to accept and adapt to changes, leading to confusion and delays. The process of change is perceived as complex and time-consuming, which adds to the challenge of implementing DevOps [2, 15].

Table 2. DevOps risks and risk mitigation strategies

DevOps risk	Risk mitigation strategies
Lack of idea about the project	Make the process visible and transparent so that people can relate with the work
Tacit knowledge is not strong	Giving priorities to the resources so that key personas feel valued, if they are working for other systems the company should find someone so that the extra load is relieved, and they can concentrate fully on their project
Sudden change in team culture	By using change management, it is possible to eliminate the work impact, Teams need to have acceptable, and some resources might not work efficiently. Teams need to cope with the existing situation
The Budget allocation for toolsets is important because wrong choices create risks for the project	Experts should be involved who are good at tool agnostics and experienced in shortlisting what could work for the environment
Lack of Communication with Developers and the clients	Better management strategies required so that everyone has a clear idea about the process
Improper code review by team members	This code review should be effective and it should be associated with the proper test

Lack of focus or differences in development is another challenge for DevOps practices. Often developers faced that there is a lack of focus in the development process. They are not sure of what they are doing, there could be miscommunication with team members. There could be misconceptions between development teams and operations team members. Due to these reasons, differences occur in the development process [1,3,15,21].

Creating proper test and production environments is a significant challenge. Both testing and production environments are crucial for the production process, and it is essential to have a well-designed testing process for the code. The production environment should support the testing process seamlessly. Poor integration can hamper the testing process, which is why it is essential to set up proper test setups to ensure that the rest of the process functions effectively [9].

Choosing the right tools for DevOps operations is another obstacle companies face. They select the tools based on their project needs and requirements. However, finding or selecting the appropriate tools is often difficult for companies.

5.2 Research Limitations

We witnessed certain limitations in conducting the research. First of all, the research did not consider the psychological aspects of the DevOps working environment and could not cover the emotional aspects of employees working in teams.

Second, In the study, practices of IT organizations were observed but the focus was developed countries IT practices. Therefore, if we could consider developing countries' DevOps operations then we could compare the scenario of developed and developing countries' IT practices to understand better views on DevOps challenges and risk mitigation strategies.

Third, due to lack of time we could not conduct longitudinal studies. A prospective study would be a great way to focus on DevOps practices which might help the researchers to understand management practices with experts' perspectives over time.

Fourth, our topic is narrowed to DevOps operations and organizational practices. Due to this reason, the domain became more specific. Identifying DevOps professionals for interviews was specifically a real challenge. We had to use various techniques to find professionals for interviews. Finding professionals was difficult and considered one of this research's major limitation, as there is a possibility of response bias and selection bias.

Fifth, the respondents could not share some information that they consider confidential for their companies. Due to those issues, we could not ask them questions as planned.

5.3 Future Research

We have identified several areas in the DevOps domain that require further study.

Performing a comparative study In the future, we will perform a comparative study that covers different IT organizations using DevOps practices. As we know different organizations have different DevOps practices and the challenge and risk mitigation factors might not be the same for all organizations. The implementation and adoption of DevOps might vary for various organizations.

Conducting longitudinal research DevOps collaboration culture is one of the core concepts for DevOps practices. We could try to focus on a longitudinal research study by observing for an extended period of time. Thus, we can get better insights and overviews of DevOps collaboration culture in organizations.

Research model for identifying risks and mitigation strategies for success factors We propose the development of a novel model that addresses DevOps challenges and incorporates critical success factors. Such a model would serve as a valuable framework for identifying and mitigating various risks within the organization. By leveraging this model, we can establish a

comprehensive understanding of the factors that contribute to success in this domain and develop effective strategies for addressing any challenges that arise.

Combining DevOps and MLOPs for better performance The incorporation of artificial intelligence (AI) within DevOps presents a promising opportunity to elevate performance to new heights. By leveraging AI in the software development life cycle, DevOps can streamline operations, resulting in more expedient development and improved operational cycle performance. This translates to a more positive user experience, as new AI features are implemented within DevOps. Moreover, the utilization of machine learning algorithms enables the collection of data from a multitude of sources, further enhancing the potential of AI and DevOps. This research area holds much promise, as it opens up new avenues for developing a diverse range of AI models within DevOps.

Developing scales for conducting survey Developing scales for measuring success and risk factors could be a great approach for doing future research. We observed that there were few studies that focused on scale development. These scales could be a great tool for quantitative surveys to collect data from professionals.

6 Conclusions

The seamless collaboration between development and operations teams, fostered by the DevOps cultural movement, is critical in streamlining the software development life cycle. Our extensive research, which included 26 semi-structured interviews with DevOps professionals, has identified numerous risk factors with mitigation strategies encountered during the implementation and adoption phase of DevOps. Our research has identified four main risk factors and several risk mitigation strategies by companies that practice DevOps. It is of utmost importance that this study guides future research agendas and delves further into the DevOps domain.

References

1. Azad, N.: Understanding devops critical success factors and organizational practices. In: 2022 IEEE/ACM International Workshop on Software-Intensive Business (IWSiB), pp. 83–90. IEEE (2022)
2. Azad, N., Hyrynsalmi, S.: What are critical success factors of devops projects? a systematic literature review. In: Wang, X., Martini, A., Nguyen-Duc, A., Stray, V. (eds.) Software Business. 12th International Conference, ICSOB 2021, Drammen, Norway, December 2–3, 2021, Proceedings, pp. 221–237. Springer (2021)
3. Bass, L., Weber, I., Zhu, L.: DevOps: A software architect’s perspective. Addison-Wesley Professional (2015)
4. Berg, J.M., Dutton, J.E., Wrzesniewski, A.: Job crafting and meaningful work (2013)

5. Bezemer, C.P., et al.: How is performance addressed in devops? In: Proceedings of the 2019 ACM/SPEC International Conference on Performance Engineering, pp. 45–50 (2019)
6. Bou Ghantous, G., Gill, A.: Devops: Concepts, practices, tools, benefits and challenges. PACIS2017 (2017)
7. Charmaz, K.: ‘discovering’ chronic illness: using grounded theory. *Social Sci. Med.* **30**(11), 1161–1172 (1990)
8. Debois, P., et al.: Devops: a software revolution in the making. *J. Inform. Technol. Manage.* **24**(8), 3–39 (2011)
9. Ebert, C., Gallardo, G., Hernantes, J., Serrano, N.: Devops. *IEEE Softw.* **33**(3), 94–100 (2016)
10. Fitzgerald, B., Stol, K.J.: Continuous software engineering: a roadmap and agenda. *J. Syst. Softw.* **123**, 176–189 (2017)
11. Gartner, I.: Gartner it glossary devops. Gartner IT Glossary (2017)
12. Gioia, D.A., Patvardhan, S.D., Hamilton, A.L., Corley, K.G.: Organizational identity formation and change. *Acad. Manag. Ann.* **7**(1), 123–193 (2013)
13. Gupta, V., Kapur, P.K., Kumar, D.: Modeling and measuring attributes influencing devops implementation in an enterprise using structural equation modeling. *Inf. Softw. Technol.* **92**, 75–91 (2017)
14. Jones, S., Noppen, J., Lettice, F.: Management challenges for devops adoption within UK smes. In: Proceedings of the 2nd International Workshop on quality-aware devops, pp. 7–11 (2016)
15. Krey, M., Kabbout, A., Osmani, L., Saliji, A.: Devops adoption: challenges & barriers. In: 55th Hawaii International Conference on System Sciences (HICSS), virtual, 3–7 January 2022, pp. 7297–7309. University of Hawai‘i at Manoa (2022)
16. Leite, L., Rocha, C., Kon, F., Milojcic, D., Meirelles, P.: A survey of devops concepts and challenges. *ACM Comput. Surv. (CSUR)* **52**(6), 1–35 (2019)
17. Lwakatare, L.E., Kuvaja, P., Oivo, M.: Dimensions of DevOps. In: Lassenius, C., Dingsøyr, T., Paasivaara, M. (eds.) *Agile Processes in Software Engineering and Extreme Programming: 16th International Conference, XP 2015, Helsinki, Finland, May 25–29, 2015, Proceedings*, pp. 212–217. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-18612-2_19
18. Macarthy, R.W., Bass, J.M.: An empirical taxonomy of devops in practice. In: 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), pp. 221–228. IEEE (2020)
19. Miles, M.B., Huberman, A.M.: *Qualitative data analysis: An expanded sourcebook*. sage (1994)
20. Nguyen-Duc, A., et al.: Work-from-home and its implication for project management, resilience and innovation - a global survey on software companies (2022). <https://arxiv.org/abs/2202.04950>
21. Riungu-Kalliosaari, L., Mäkinen, S., Lwakatare, L.E., Tiihonen, J., Männistö, T.: DevOps adoption benefits and challenges in practice: a case study. In: Abrahamsson, P., Jedlitschka, A., Nguyen Duc, A., Felderer, M., Amasaki, S., Mikkonen, T. (eds.) *Product-Focused Software Process Improvement*, pp. 590–597. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-49094-6_44
22. Strauss, A., Corbin, J.: *Basics of qualitative research techniques* (1998)
23. Velasquez, N.F., Kim, G., Kersten, N., Humble, J.: *State of devops report* (2014)




Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Positive Customer Experience is Enhanced by Effective Agile Practices

Riina Piironen^(✉) , Ilkka Jormanainen , and Markku Tukiainen 

University of Eastern Finland, 80110 Joensuu, Finland
riinapii@student.uef.fi, {ilkka.jormanainen,
markku.tukiainen}@uef.fi

Abstract. This paper explores the connection between agile methods and digital customer experience, aiming to identify what are the hallmarks of a good agile way of working. The research is an exploratory case study consisting of interviews and analysis. In summary, the research suggests that the hallmarks of a good agile way of working are 1) breaking down tasks into sufficiently small pieces, 2) defining tasks precisely and releasing them to production evenly, 3) continuous improvement, and 4) good planning of sprints. These good agile operating methods can be seen in the development measures as a short lead time, a short time to export to production, low errors, and a high deployment frequency. According to the findings, these metrics are linked to the Net Promoter Score (NPS), a measure of customer experience. A team with sufficient technical capabilities team that utilizes agile operating methods is able to produce the desired things for customers at exactly the right time while constantly improving, so that the NPS is positive, and its direction is improving. On the other hand, the team's bad operating methods are also visible in the NPS meter – in this case, the NPS fluctuates strongly. Teams can obtain insightful supplementary data about their own practices by keeping track of development measures.

Keywords: agile methods · project management · software development · agile organization · customer experience

1 Introduction

Agile methods are a set of different lightweight and quickly responsive methods and their tools. Agile methods share similar values and principles based on the agile software declaration, which helps to optimize project management [4]. For example, Scrum, Lean and DevOps are examples of agile methods.

These methods, often the challengers of the traditional process models, have grown in popularity as part of project management and goal-oriented management around the world, both in the IT sector and outside of the IT sector. They are marketed as methods for increasing customer satisfaction and the success rate and efficiency of projects [1, 2]. However, it is not entirely clear which customer experience measures show the benefits of agile methods. It is also not clear which agile way of working methods affect customer experience, the success rate of projects and efficiency.

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 386–399, 2024.

https://doi.org/10.1007/978-3-031-53227-6_27

The purpose of this paper is to explore a connection between agile methods and digital customer experience. The connection is explored through thematic interviews and analysis. The representatives of the theme interviews were selected from seven different self-directed technical teams ($n = 7$). Every team, which holds a significant role in the target organization for developing interactive mobile services, were included into the study. In this research, the following research questions are answered:

RQ1: How does an agile way of working and the technical ability supporting it affect the digital customer experience?

RQ2: In which customer experience metrics and agile metrics, we can see benefits of agile way of working?

RQ3: What are the hallmarks of good agile way of working and team's technical abilities?

To address the research questions, we have collected data in three phases. Initially an open interview was held with the target organization's goal-oriented management expert, where it was explained how the organization aims to influence the digital customer experience with agile methods. Based on the interview, themes were formed, and these themes were used to guide thematic interviews that were held with representatives of seven different teams. Customer experience and agile measures data was also collected from the organization's databases. The results of the interviews were used as explanatory factors in the analysis, which utilized data from customer experience and agile measures.

The results of this explorative case study indicate that there is a connection between agile methods and digital customer experience. The results can help teams to identify the best agile way of working methods in terms of customer experience.

2 Related Work

2.1 Agile Methods and Agile Measures

Agile methods are a set of different lightweight and quickly responsive methods and their tools. Agile methods share similar values and principles based on the agile software declaration. Agile methods such as Scrum, Lean and DevOps helps organizations to optimize their project management practices [4].

Scrum is a framework based on empiricism, i.e., experience thinking, which focuses on producing a software project that meets the customer's needs through phasing and continuous control [14, 15]. Project transparency, review, and adaptation are essentials in empirical process management, and these form the basis of Scrum.

DevOps can be considered a method of operation, whose purpose is to integrate software development and operations by narrowing the silo that traditionally exists between them [6]. DevOps can be considered as a logical extension of other agile methods such as Scrum. Software development plays an important role in DevOps automation, customer orientation and operational transparency [8].

In a key role in providing services and delivery is an agile self-directed team. The faster the team is able to make changes to the services, the faster customers can be offered value, and more likely, the customer experience will be positive. It is important to measure the performance of a team that uses agile methods in order to verify possible problem

areas in developing of services and thus increase the team's performance [5, 15]. The DevOps Research & Assessment (DORA) team has identified five key agile measures that can be used to measure development team's performance. The measures are the following: lead time for code changes, time to restore service, deployment frequency, change failure rate and reliability. With the help of these metrics, teams can be classified as top-level teams or low-level teams [5]. For example, a team with a short lead time is typically at the top level. The target organization of the research uses the same measures that DORA team identified, and the teams involved in this research were selected using these measures. There are teams that are at the top level in light of these measures and there are teams that are at a low level.

2.2 Digital Customer Experience and NPS

Digital customer experience can be defined as the customer's internal and subjective reaction to a digital product or service the customer interacts with [16]. The digital customer experience consists of all the organization's offering-quality, customer service, advertising, product and service features, usability and reliability of the product or service affect the customer experience [9]. The most important characteristics of a digital service in terms of digital customer experience are speed, functionality, performance, ease of use and reliability, as well as minimal errors [7]. In an ideal situation, product developers know how to develop a product forward based on how customers use the products or services and which issues in the product frustrate customers [9].

Customer experience can be measured with, for example, the NPS (Net Promoter Score) meter. NPS measures the customer's willingness to recommend, i.e., whether the customer would promote the organization or its services to others. NPS boils down to the question "On a scale of 0–10, how likely would you recommend our services/products to a friend or a family member?" Based on the points given, customers can be divided into promoters, passives, and detractors, that is, customers who are dissatisfied with the service. NPS is calculated using the formula $\%Promoters - \%Detractors = NPS$ [16].

2.3 Connection Between Agile Methods and Digital Customer Experience

The connection between agile methods and digital customer experience has not been studied at a sufficient level. There are only a few research papers discussing this topic.

According to Aghina et al. [1, 2], customer satisfaction can be improved by up to 30 percent with the help of agile methods. However, the report does not reveal which agile way of working methods affect the customer experience.

Bambauer-Sachse and Helbling [3] have studied the connection between agile methods and customer experience in a B2B context. In the B2B context, according to authors, satisfaction with the process is a more important factor in general customer satisfaction than satisfaction with the end result of the service [3]. Thus, Bambauer-Sachse and Helbling [3] look at the issue from a different perspective as we do in our case study, where the customers are not companies, but end users of a digital product.

According to Recker et al. [11], agile methods have a positive effect not only on the customer experience but also on the product's functionality, quality and staying within the budget. The research does not so much take a position as to which agile way of

working methods affect these positive results – instead, according to authors, different development practices influence the outcome. So, even this paper by Recker et [11] is written from different perspective than this paper.

According to Olteanu [10], projects are completed faster and with less bugs with the help of agile methods compared to the traditional waterfall model. The research states that agile methods have influence the customer relationship but does not elaborate the more detailed effects.

As evident from the above, there is noticeable void in the current literature, as no studies address the exact extent to which team's agile practices can influence the final customer experience. The results of this paper take one step towards a more comprehensive understanding related to the topic.

3 Research Approach

The objective of the case study presented in this paper is to embody the connection between agile methods and digital customer experience. Three research questions have been set for the research. These questions are answered with the help of a case study. The focus of the research is on an organization that creates mobile services with interactive features. These services are utilized by hundreds of thousands of individuals. By “customers” in this paper, we mean end-users who use these mobile services. The target organization's most important customer experience measure is NPS. In the target organization, NPS can be anything on a scale of -100 to 100.

To answer the research questions, we have used a process consisting of three steps (Fig. 1). In the first stage, an open interview was held with the target organization's target management expert. In interview, it was mapped out how the organization strives to influence the digital customer experience with agile methods. Based on the interview, themes were formed, which were used later to guide the thematic interviews.

Before phase two (theme interviews), we had to identify the teams from the organization that develop these interactive mobile services, and whose customers are end-users. The target organization reports the performance of the teams considering different agile measures. Some teams are at the top level in the measures, there are mid-level teams, and teams at a lower level. We identified seven teams suitable for the research. These seven teams develop interactive mobile services for end users in the organization. Two teams are at the top level and five at the low level in terms of agile measures. All teams have nine developers and a product owner. The teams are therefore similar in composition.

In phase two (theme interviews) we interviewed representatives of all seven teams. The representatives were the product owners of the teams. In the target organization, the product owner is responsible for maximizing the value of the product and the work of the development team, and the practical tasks include managing the product's development queue and communicating with different stakeholders. Finally, in phase three (analysis), the data collected from the interviews were used as explanatory factors for analysis that used digital customer experience measures and agile measures.

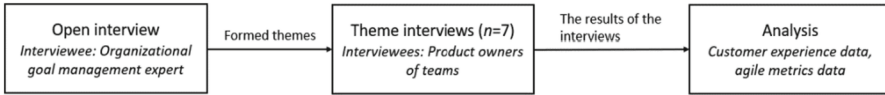


Fig. 1. Phases of the research

3.1 Data Collection

Data was collected from the previously mentioned interviews, which were eight in total – one open interview and seven thematic interviews. With the help of an open interview, data was collected on how the organization aims to influence the digital customer experience using agile methods. The purpose of the theme interview was to express and collect data on how the themes extracted from the open interview guide the team’s activities and to look for hallmarks of a good agile way of working. All interviews were conducted remotely, and each interviewee gave consent for the answers to be used for research purposes. However, the answers are processed in such a way that the identity of the respondent (or the team) is not identifiable.

In addition to the interviews, data was collected from the organization’s databases. For analysis, the data has been aggregated to the monthly level.

3.2 Data Analysis

The first phase of the analysis was the transcription of the open interview. After transcribing the open interview, the material was divided into themes, which is one of the work phases of qualitative analysis [13]. The material was divided into themes in Word by color-coding the written material so that sentences related to the same theme were marked with the same color. One researcher worked through the material in three rounds of iteration, re-color-coding the sentences and checking if they were still classified under the same topic. The data collection and classification were originally done as a thesis work of the main author of this article. Thematization can be considered an interpretative act [12], and in this research thematization requires subjective interpretation due to the nature of interview. Therefore, only one researcher has been involved in the thematization of the material, but the thematization was discussed with the supervisor of the work.

In the end, it was settled on the following themes: self-direction, common goals, continuous learning, continuous improvement, the ability to understand the needs of customers and the ability to get things done. These are the themes with which the organization strives towards a better customer experience. For example, the sentence.

“And refactoring is the choice of this model. Because we work iteratively, we are constantly in a situation where we have to build the same thing again”
(Organizational project management expert).

was classified under the theme continuous improvement. The sentence.

“At the same time, we learn all the time and are able to focus what we do in even smaller pieces more precisely on the goal”
(Organizational project management expert).

was classified under the theme continuous learning. The theme of the ability to get things done was classified as, for example, the sentence:

“The work must be done in order for it to produce any value for the customer”.
(Organizational project management expert).

When the theme interviews were also transcribed, we started looking for connections in the collected data. We look for a connection between Agile measures and NPS data by doing a cross-comparison by gathering all the teams in the same table. One table dealt with the emergence of agile way of working methods, agile measures, production usability, recovery from disruptions and customer satisfaction by classifying these into levels: low, average, good, high. In another table, we compiled the differences and similarities between the teams. The table covered agile measures, monitoring customer feedback, continuous learning, continuous improvement, shared goals, release to production cycle and the ability to complete sprint tasks. Finally, we started looking for explanatory power for the observations in the tables in the materials of the thematic interviews. The results are presented in the next section.

4 Results

In this section, we present the results of our case study. Based on the analysis, it is possible to identify the connection between agile methods and customer experience. Based on the results, it is also possible to compile the best practices for improving the customer experience using agile methods.

4.1 Teams at the Top Level in the Light of Agile Measures

According to the agile measures, the top-level development teams unfortunately did not fully fit in the scope of the research, as the customers are internal customers and not actual end-users. However, it is still important to address the interview results of these teams to gather the best practices that make these teams top teams. Let's call these teams A and B. The teams utilize Scrum and DevOps.

The common goals are reflected in the prioritization of the team's tasks and in directing the activities. Agreed goals are given high priority. Self-directedness is perceived as the freedom to decide on the team's ways of working and to make decisions independently.

Continuous learning is always done in teams as needed. Team B uses shared learning. One member studies a new thing. After this, the team member goes through the new issue with the rest of the team, teaching and supporting others. Team B feels that they have sufficient technical ability to solve various problems.

The teams consider continuous improvement in their operations. Technical debt is dismantled in teams by refactoring and developing new, more sustainable solutions. Time is reserved for refactoring in sprints.

In team B, the sprints are planned so that 60% of the working time is reserved for tasks that are known in advance. The remaining 40% of the working time is reserved for tasks that cannot be predicted in advance. The tasks are broken down into small enough entities so that it is possible to implement them during the two-week sprint. Each task is also defined precisely enough, and not a single task is taken up until it has been defined precisely enough.

Team B's agile way of working methods support development efficiency. According to the interview results, it is essential to maintain the team's skills and to plan the sprints accurately. When planning sprints, one should take into account 1) the available working time 2) things that cannot be prepared for in advance 3) splitting the tasks into sufficiently small entities, and 4) that everyone's task should be defined sufficiently precisely, before it is taken to the agenda.

4.2 Teams at the Low Level in the Light of Agile Measures

Five low-level teams participated in the research. However, three teams were dropped during the analyze when it was found that customer experience data (NPS) for the observation period was incomplete. Hence, this study includes the teams identified as D and E. Both teams work using Scrum. Work is controlled in teams both with the help of a Kanban board and also with product and sprint backlogs. Both teams have also features of DevOps – work is done in a customer-oriented manner with continuous improvement, software development is aimed to be automated as far as possible, and the service of each team is monitored.

Table 1 shows that the teams perform similarly to each other. The following scale is used in the table: high, good, average, low. For both teams, the measurement of customer satisfaction is increasing. In both teams, agile way of working methods is manifested at some level, but every team has a lot of room for improvement – tasks should be broken down into smaller ones, tasks should be defined more precisely, and release pipelines could be automated more. Both teams' service uptime has been 99.7% to 100% during the review period, meaning that the service has been available to customers 99,7% - 100% of the time. The generally targeted service uptime is 99% [1]. The teams are able to restore the service from disruptions to a normal state quickly.

Table 1. Cross-tabulation of teams

Team	Agile measures performance	Manifestation of agile ways of working	Service uptime	Recovery from disturbances	Customer Satisfaction
D	average	average	high	high	low, increasing
E	good	average	high	high	good, increasing

Results of the Interviews by Theme

Self-directedness. Teams have annual goals, quarterly goals and sprint goals. If necessary, changes can be made to plans and priorities even with a fast schedule - for example, critical production errors always come before planned issues.

Agile measures. Teams are familiar with Agile measures, but they do not guide the teams' activities. Teams have the ability to make a production release whenever the

defined quality criteria are met. According to agile principles, agility is ability to put code into production every day, but only make release visible to customers as needed.

Ability to understand customer needs. Customer feedback is actively monitored, and based on customer feedback, a lot of work is added to the teams' to-do lists. The teams perform customer testing if necessary. Errors reported by the customers will be corrected immediately. The teams also have real-time monitoring of their service.

Continuous learning and continuous improvement. Work time is set aside for continuous learning. New things are often learned while doing work. Teams hold retrospectives. Teams are also working to eliminate technical debt.

Ability to get things done in a sprint. Team D defines the tasks precisely - every task has a definition of done. However, the tasks are not broken down into small entities, because the team sees that it takes a lot of working time - because of this, the planned tasks are not always completed. In team E, the definition of done is defined for the tasks. The team tries to take on only tasks that can be completed during the sprint. No implementation and testing of the feature, however, is never done in the same sprint, so the set of tasks is also not completed during the same sprint.

4.3 The Connection of Agile Measures to Customer Experience Measures

Team D

Table 2 presents the key figures of the meters every six months. As Table 2 illustrates, when the development measures are low, customer experience is also low. When the indicator values are increasing in the second half of the year, also the customer experience has turned to growth at the same time.

Table 2. Team D's measures development

	Lead time	Export to production lead time	Deployment frequency	Customer experience (-100 to 100)
The first half of the year	low, 7 months (1 – 6 months)	low, 6 months (1 – 6 months)	low (once a month – once every 6 months)	low
The second half of the year	low, 2 months (1 – 6 months)	low, 1 months (1 – 6 months)	average (once a week – once a month)	low, increased by ~ 50 units

In the first half of the year, development was done on the previous application platform, which had deteriorated a lot in terms of quality, so development and release to production was extremely slow. With the new application platform, architecture and user interface, development had become easier and faster - it can be seen in the team's development measures as a positive development in the second half of the year. Both the old and the new application platforms contain all the backend and frontend features needed by an interactive mobile service.

Customer feedback provides a lot of work for teams' backlogs. As the team's performance increases, the team can complete new development tasks and bug fixes faster and

take them to production faster than before. When the customer’s needs are met faster, the customer experience also seems to improve. The improvement of the customer experience is also influenced by the new user interface developed by the team, for which customer testing was carried out.

The team’s performance would increase even more if the team reserved, for example, 40% of the working time of the sprints for unexpected tasks, such as production errors, and split the tasks into smaller entities. In this case, the tasks would be completed faster, which would reduce the lead time. The team has the ability to release to production whenever various quality criteria are met, so as the lead time decreases, new features and bug fixes could also be released to production faster and more often. Hypothetically, it is entirely possible that the team’s efficiency and customer experience would improve even more if features corresponding to the customer’s needs could be released more often into production. In the case of Team D, however, it can be stated that agile way of working enables the team’s performance efficiency, which would seem to improve the customer experience.

Team E

Table 3 presents the key figures of the indicators every six months for team E. It can be seen from the table that the indicators of development have not developed significantly in a positive or negative direction. Is it remarkable that customer experience fluctuates by twenty units every quarter in both negative and positive directions. However, the interview material **did not provide explanation** for growth or fluctuations in customer experience, so we explored further some external factors not mentioned in the interviews. We started looking for explanatory power by listing things that affect the customer experience and excluding options one by one.

Table 3. Team E’s measures development

	Lead time	Export to production lead time	Deployment frequency	Customer experience (-100 to 100)
The first half of the year	low, 2 months (1 – 6 months)	average, 0,5 months (1week – 1 months)	average (once a week – once a month)	good, fluctuates quarterly
The second half of the year	low, 3 months (1 – 6 months)	average, 0,7 months (1week – 1 months)	average (once a week – once a month)	good, increased by ~ 5 units

Seasonal Variation. First, it was investigated whether the service is related to a possible seasonal variation in the customer experience. This would be reflected in the fact that each year similar trends in the customer experience would be found around the same time. The alternative was investigated by comparing four years of customer experience data. Customer experience fluctuates by twenty units every year, but the moments of fluctuate are not the same yearly. The increase or decrease in the customer experience is therefore not caused by seasonal changes.

Increased Volume in Interactive Mobile Services. As another option, we investigated whether, for example, there could be more volume in the summer than at other times, when the processing times would be longer and this would be reflected in the customer experience as a ripple. However, this option was ruled out, because the customer experience meter in this case does not measure the customer experience from the beginning to the end of the process. So the duration of the processing times does not affect the customer experience, because the customer answers the survey before knowing how long the processing will take.

Digital Service Performance. As a third option, an attempt was made to find out whether there have been changes in the performance of the service, which would appear as a decrease in the customer experience. However, the service's uptime is 99.7%, so this is an unlikely option.

Features Published for Production. The team publishes large releases that contain many different features. These big releases are made quarterly. Production errors also fluctuate quarterly. The number of errors seems to increase in the next quarter after a big release has been put into production. With big releases, the number of production errors increases. When we reflect releases and errors in the customer experience, we notice that as production errors increase, the customer experience deteriorates. As the number of production errors decreases, the customer experience improves. Figure 2 illustrates this phenomenon.

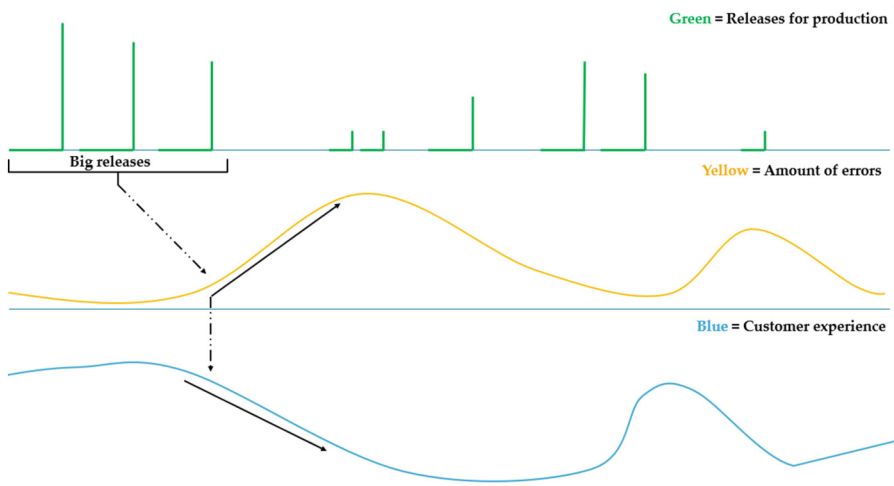


Fig. 2. Connection of the number of releases and errors to the customer experience

The working methods of the development team are the root cause of the fluctuations and improvements in customer experience throughout the year. The development team does release-driven work, i.e., releases larger entities for production at once. The way of working is reflected in production: errors, and customer experience fluctuates. Customer experience fluctuates in both negative and positive waves. Negative waves are seen when

the team has released large releases. Positive waves are seen after the team has fixed the bugs and errors in production.

Good agile way of working methods can be seen in the development measures as a short lead time, a short export to production time, low error rates and a high deployment frequency. Based on the case study, these measures have a connection to the NPS measure of the customer experience. A technically capable self-directed team is able to produce the desired things for customers at exactly the right time while constantly improving, in which case the NPS is positive and in an improving direction. Bad working methods of the team are also visible in the NPS meter - in this case, the NPS fluctuates strongly.

Based on the interviews, the development team could release to production whenever the quality criteria are met - for one reason or another, however, they do not use this ability. Furthermore, the development team never carries out feature implementation and testing during the same sprint – this is not in line with agile ways of working, as this causes the lead time to increase.

If the team used their ability to release to production every time the quality criteria are met and did the implementation and testing of the feature during the same sprint, the lead time and export time to production would be shortened. In addition, with a steady pace of releases to production, potential errors would be distributed more evenly, and they could be corrected more efficiently - there would not be so strong fluctuation in customer experience. The increase in customer experience during the second half of the year is probably not due to the efficiency of the team's work, but due to the fact that the development team has corrected errors in production.

5 Discussion

5.1 Key Findings

The purpose of the research is to demonstrate the connection between agile methods and digital customer experience. Based on this research, it can be suggested that the implementation of agile methods appeared to have a positive impact on customer experience. However, further research is needed to confirm this assertion.

RQ1: How does an agile way of working and the technical ability supporting it affect the digital customer experience?

When the tasks are precisely defined and broken down into small enough pieces, they can be completed faster, which reduces the lead time, and the team has an opportunity to release to production more often. If this option is used, the deployment frequency of the team will also improve. These enable the customer's needs to be met more efficiently and thus improve the customer experience. In addition to being efficient and technically capable, the teams must be able to take into account the customer's needs and react to them, as well as be able to quickly correct possible production errors.

RQ2: In which customer experience and agile metrics, we can see benefits of agile way of working?

Good agile way of working methods can be seen in the development measures as a short lead time, a short export to production time, low error rates and a high deployment frequency. Based on the case study, these measures have a connection to the NPS measure

of the customer experience: a technically capable self-directed team is able to produce the desired things for customers at exactly the right time while constantly improving, in which case the NPS is positive and in an improving direction. Bad working methods of the team are also visible in the NPS meter - in this case, the NPS fluctuates strongly.

RQ3: What are the hallmarks of good agile way of working and team's technical abilities?

Hallmarks of a good agile way of working are breaking down tasks into small enough pieces, defining tasks precisely and releasing them to production evenly, continuous improvement and good planning of sprints. These hallmarks are best practices as well. When planning a sprint, one should also consider things that cannot be prepared for in advance by reserving, for example, 40% of the sprint's working time for unexpected things. In addition to agility, the team must also be technically capable so that the team can produce a high-quality and reliable service or product for the customer.

The Importance of Agile Measures

Teams could have paid more attention to agile measures, as they can provide valuable additional information about team operations. A long lead time can indicate that the task sets are too large. A low deployment frequency can indicate that the team is not using its ability to release features and bug fixes to production optimally.

5.2 Limitations

A limitation of the research is the small sample size ($n = 7$). However, in this case all the teams that play key roles in the target organization in developing interactive mobile services were included in the research. The final amount of analyzed ($n = 4$) teams was also small, since otherwise suitable teams had to be dropped from the study due to the lack of customer experience data. With lack of customer experience data, it would have been impossible to make a reliable analysis.

The NPS metric is not designed to provide actionable insights into problems in digital customer experience [16]. To get more detailed information about different problems, other metrics are needed for support.

Another limitation is that the teams in this research work in a narrow sector. Thus, the generalizability of the results to other sectors is not guaranteed without further evaluation.

6 Future Research

With the help of the findings of the research, topics were found that require further research. These topics can significantly improve the optimization of agile methods.

Before and After Optimization Using Best Practices

In the future, the connection between agile methods and customer experience could be studied in more detail over a longer period of time. It would be meaningful to include a period before in the study optimizing and post-optimizing agile team practices. After this, the customer experience could be more closely reflected in the team's operating methods and agile measures.

Optimizing the Size of the Release from the Point of View of Customer Experience

Our results show that releases that are too large lead to more errors, resulting in decreased customer experience. It would be important to study what is the optimal publication size so that it affects the customer experience positively. The research should also identify the effects of too large releases.

Taking Open Customer Feedback into Account

Open customer feedback could be used in future research. The research could analyze how the customer experience develops when the team implements the wishes and needs expressed in open customer feedback.

Replicating the Research on a Larger Scale

Replication of the research would bring significant value to the software industry. The research would be done on a larger scale, so the results of the research can be generalized. In addition to NPS, the research would also use other customer experience metrics, such as CES and FCR.

7 Conclusions

Agile methods are widely used around the world. They help development teams work efficiently and react to changes quickly. Optimizing agile methods could help organizations improve customer satisfaction continuously. Optimization should always start by looking at the numbers of agile measures and analyzing the reasons for those numbers. Based on this research, the best practices from the point of view of agility have been listed, which help to improve the customer experience. They are as follows: breaking down tasks into sufficiently small ones into pieces, precise definition of tasks and steady release to production, continuous improvement and good planning of sprints.

References

1. Aghina, W., Handscomb, C., Salo, O., Thaker, S.: *The Impact of Agility: How to Shape Your Organization to Compete*. McKinsey & Company (2021)
2. Aghina, W., Handscomb, C., Ludolph, J., Róna, D., West, D.: *Enterprise Agility: Buzz or Business Impact?* McKinsey & Company (2021)
3. Bambauer-Sachse, S., Helbling, T.: Customer satisfaction with business services: is agile better? *J. Bus. Ind. Market.* **36**(8), 1389–1402 (2021)
4. Beck, K., et al.: *Manifesto for Agile Software Development* (2001)
5. DeBellis, D., Peters, C.: *2022 Accelerate State of DevOps Report*. Google/DORA (2022)
6. Ebert, C., Gallardo, G., Hernantes, J., Serrano, N.: *DevOps*. *IEEE Softw.* **33**(3), 94–100 (2016)
7. Grisham, P.S., Perry, D.E.: Customer relationships and extreme programming. *ACM SIGSOFT Softw. Eng. Notes* **30**(4), 1–6 (2005)
8. Kim, G., Humble, J., Debois, P., Willis, J., Forsgren, N.: *The DevOps Handbook: How to Create World-Class Agility, Reliability, & Security in Technology Organizations* (Second). IT Revolution Press (2021)
9. Meyer, C., Schwager, A.: Understanding customer experience. *Harv. Bus. Rev.* **85**(2), 116–126 (2007)

10. Olteanu, C.G.: IT Agile Transformation. *Econ. Inf.* **18**(1), 23–31 (2018)
11. Recker, J., Holten, R., Hummel, M., Rosenkranz, C.: How agile practices impact customer responsiveness and development success: a field study. *Proj. Manag. J.* **48**(2), 99–121 (2017)
12. Saldana, J.: *The Coding Manual for Qualitative Researchers*. Sage, London (2009)
13. Sarajärvi, A., Tuomi, J.: *Qualitative research and content analysis*. Tammi (2018)
14. Schwaber, K., Beedle, M.: *Agile Software Development with Scrum*. Upper Saddle River: Prentice-Hall, New Jersey (2001)
15. Srivastava, S., Trehan, K., Wagle, D., Wang, J.: *Developer Velocity: How Software Excellence Fuels Business Performance*. McKinsey & Company (2020)
16. Wiedenhofer, L.: The Importance of Acting Today: Getting digital experiences right is not optional. In: Wiedenhofer, L. (ed.) *Digital Customer Experience Engineering: Strategies for Creating Effective Digital Experiences*, pp. 1–3. Apress, Berkeley, CA (2021). https://doi.org/10.1007/978-1-4842-7243-5_1



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Information-Centric Adoption and Use of Standard Compliant DevSecOps for Operational Technology: From Experience to Design Principles

Henry Haverinen¹ , Tero Päivärinta²  , Jussi Vänskä³ ,
and Henry Joutsijoki⁴ 

¹ Cyberismo Oy, Saarikonkuja 26, 37500 Lempäälä, Finland
henry.haverinen@cyberismo.com

² Faculty of Information Technology and Electrical Engineering, M3S, University of Oulu,
Oulu, Finland

tero.paivarinta@oulu.fi

³ Valmet Automation Oy, Lentokentänkatu 11, 33900 Tampere, Finland
jussi.vanska@valmet.com

⁴ Insta Automation Oy, Sarankulmankatu 20, 33900 Tampere, Finland
henry.joutsijoki@insta.fi

Abstract. Secure and agile development of operational technology (OT) and related software in industry is a crucial but challenging issue. Generally recognized standards such as IEC 62443–4–1 set up the requirements for cybersecurity processes for OT and software development. The main challenge of IEC 62443–4–1 resides in its adoption and implementation in practice, which originates from the standard’s complexity. We propose three novel design principles and two subsequent design objectives to be prioritized for future design-research oriented work on standard-compliant DevSecOps. The design principles have been formed after six years of experience and observations in cybersecurity consulting in industry, documented here as a piece of action design research (ADR). As a case study, we describe instantiation of the design principles at Valmet Automation Systems, one of the earliest IEC 62443–4–1 -certified companies. The proposed design principles altogether suggest for the information-centric view on the contextual adoption and use of the IEC 62443–4–1 standard in DevSecOps practices for OT.

Keywords: DevSecOps · operational technology · IEC 62443–4–1 · design principle · action design research · information-centric adoption

1 Introduction

DevSecOps is an emerging approach to software development denoting integrated security controls and practices, and security teams, throughout the tasks of the development and system operations (DevOps) cycle [13, 14]. While the agile combination of development and operations as such was introduced more than a decade ago [6, 9], the integration

challenge of security issues into agile development has continued in practice [15] and research [1, 14] alike. Reviews by Rajapakse et al. [14] and Akbar et al. [1] have outlined a great many challenges, 21 and 18, respectively, for DevSecOps adoption and management. While the industrial domain requires well-synchronized DevOps of software together with operational technologies (OT), the challenge of implementing secure coding standards, testing for security in DevOps and the sheer knowledge of role of security in connection to system and software development remain as the prioritized problem areas in the software industry [1].

In industries that involve cyber-physical systems, such as automation and control systems, information technology (IT) and OT need to be converged [4]. Such systems rely on the use and adoption of standards. The IEC 62443–4-1 standard focuses on cybersecurity during the development lifecycle especially in automation and control systems [7]. Although the standards in general form a basis to secure software development, their adoption, implementation, and operationalizing in practice is a time-consuming and laborious process. One of the core challenges of adopting DevSecOps for OT and related software is the very adoption of the often-complex security standards, such as IEC 62443–4-1, so that the professionals would also be able to operationalize the standard requirements in the development process synchronized with operations. Several issues related to this challenge are highlighted in [1, 10, 12, 14] but the empirical research on adoption and implementation of standards (e.g., IEC 62443–4-1), with actual software processes and tools, is still in its infancy [1].

Among the earliest research efforts on adoption of IEC 62443–4-1 in agile development of industrial systems, Moyon et al. [11, 12] suggest process models to be used collaboratively by security and development professionals to reach a common understanding on how to operationalize standard compliant DevSecOps. They [12] suggest that process/task-oriented understanding of the standard, indeed, becomes easier after modelling the resulting practices in the process form (with a business process modelling notation). While Moyon et al. [11, 12] provide, to our knowledge, the first demonstrations of the potential usefulness of their suggested approach, they do not report how and whether their process-based view has been operationalized in practice. Room for additional research on the security standard adoption challenge in connection to agile software development thus exists. Keeping this in mind, our research provides an early report on longitudinal experiences of actual adoption and consulting process for operationalizing IEC 62443–4-1 in the DevSecOps context of industrial automation systems.

Our research set out with a research question: How to operationalize the requirements of IEC 62443–4-1 security standards in agile DevOps of industrial automation systems? Our action design research (ADR) [16] effort covers four years of consultation and collaborative development for support practices and tools for standard adoption. Insta (<https://www.insta.fi/en/en/>) is a security consulting company working both in-depth and longitudinally with several customer organizations and cases simultaneously. In this research, Insta had an interest in developing practices and tool support for standard compliant DevSecOps. A main contribution to such formalized experience comes from Valmet Automation Systems (VAS), which is an early certified adopter of the IEC 62443–4-1 standard with its certified ISASecure® [8] SDLA (Security Development Lifecycle

Assurance) process. VAS is a business line within Valmet corporation (<https://www.valmet.com/automation/>), which has co-operated with Insta over several years. This process included researchers from both Valmet and Insta, as well as a researcher from academia. The contributions of the paper can be summarized as follows:

1. Proposing three design principles for adopting and implementing IEC 62443–4-1 standard in practice.
2. Proposing two design objectives regarding portable process information model representation and process information model management tool.
3. Presenting a real-world case by Valmet Automation Systems where the DevSecOps process has been adopted and operationalized.

2 Methodology

The ADR [16] method focuses on co-operation between researchers and practitioners to create new knowledge. The ADR approach denotes that relevant research on IT artefacts benefits greatly from collaboration with advanced organizations developing, adopting, and utilizing the innovative artefacts in question [16]. The ADR process usually takes place in iterations over time with the stages of:

1. Problem formulation.
2. Building, intervention, and evaluation (BIE), usually in multiple cycles.
3. Reflection and learning.
4. Formalization of learning and outcomes [16].

The reported ADR process covers the time frame from 2016–2022, focusing mainly on Insta-Valmet co-operation, complemented with eventual other relevant consulting experiences by Insta of the subject matter.

2.1 Two Development Cycles: 2016–2019

Prior to this research during the 2010's, such commercial concepts as BSIMM (Building Security in Maturity Model), OpenSAMM (Open Software Assurance Maturity Model), and OWASP (Open Worldwide Application Security Project) were discussed among the practitioners in Insta and VAS alike. These concepts focus on assessing the maturity and planning the adoption roadmap on a high-level, while providing limited practical support for adoption in R&D teams and no real-time visibility to adoption status. The first development cycle started in 2016. The goal was to develop an improved DevSecOps framework for VAS at R&D team-level and certified to comply with IEC 62443–4-1. In hindsight, this first iteration already resulted in several lessons towards the information-centric approach.

At first, the goal was simply how to get DevSecOps efficiently adopted in VAS. In 2018, after years of cybersecurity and DevSecOps consulting, the practitioner authors identified a more focused question: what kind of practice(s) would speed up the adoption of standard compliant DevSecOps among industrial suppliers while being repeatable and scalable so that new persons and teams can quickly learn to apply the practices.

After a successful audit in 2019, the practitioners noticed that the metrics and mechanism in terms of adoption status as well as model's modularity and flexibility required improvement. This was the motivation for the second cycle in 2019, when the DevSecOps framework at VAS was refactored to be more modular and independent from the contextual needs of the initial R&D projects, and introduced the concept of an Internal Control, to make the adoption of the model measurable. The new model was evaluated to be successful in providing real-time visibility to the adoption status. During the evaluation at Insta, the practitioners identified an opportunity to build a separate tool that would make it easier to adopt DevSecOps in different organizations that might use different software engineering tools.

2.2 The Third, Fourth and Fifth BIE Cycles: The OXILATE Project 2020–2022

The third BIE cycle took place in from December 2019 to December 2020 when the OXILATE project started (<https://itea4.org/project/oxilate.html>) and a representative of a research organization joined the team. The development cycles from now on followed the ADR guidelines more consciously. Insta implemented a prototype of a dedicated “Dependability Tool” for managing the DevSecOps information model. During the Insta's internal evaluation phase and with Insta's customers, we (all the authors) realized that while a dedicated tool enabled improved automation and more convenient workflows, moving the management of the DevSecOps information model to a new separate tool may be challenging to adopt in practice.

In the fourth BIE cycle during the first half of 2021, Insta gathered information for “pivoting” the DevSecOps model of the third cycle and interviewed their current and potential customers about the business goals, challenges, and solutions in DevSecOps and Cybersecurity Management System adoption. The design principles and the two proposed design objectives presented in this paper are based on the evaluation of the fourth BIE cycle, and the fifth BIE cycle, which consisted of further development of VAS' DevSecOps framework (and, at the same time, Insta's reference framework) that took place during the second half of 2021 and the first half of 2022. This further development was motivated by retrospectives and end-user feedback, where we identified concrete improvement areas to simplify and clarify the information model. Formalization of learning through design principles.

The data documented in the consulting and BIE cycles consists of feedback from external auditors, meeting notes from retrospectives, formally documented continuous improvement reviews, and documentation of Insta's customer interviews. The verbal interactions and sparring between Insta and VAS practitioners, and with Insta's other customers over the years have also accumulated insight. This data has now been conceptualized as design principles and design objectives of this paper.

Sein et al. [16] suggest that the learnings from BIEs should be ultimately formalized as design principles, based on the accumulated experiences. Gregor et al. [5] suggested the generic form and components of design principles to include descriptions of implementers, their aims, the intended users, context, mechanisms, enactors, and rationales. Hence, our formalization of learning takes place through such descriptions of design principles (and design objectives for the emerging issues in the end of the last BIE).

3 Proposed Solution

Compliance with the IEC 62443–4-1 standard requires a documented development process with evidence of practicing the process. An IEC 62443–4-1 requirement generally begins with the phrase “A process shall be employed...”, after which the requirements state what must be done or what must not be done. In other words, the standard focuses on the tasks or procedures that the product supplier organization must employ.

The standard focuses on the actions that must be performed. However, it largely ignores the information artifacts that are related to the process, except as evidence to demonstrate that the processes have been practiced. The starting point for the proposed solution is the realization that the information artifacts, or the information model, also deserves attention: it is easier to understand a process if you consider both the procedures and the information model of the process, i.e., the conceptualization of information used and produced. The realization about the importance of the information model resembles Fred Brooks’ [2] famous remark about the relationship between code and data structures. We have modified the Brooks’ quote to support the proposed solution as follows: “*Show me your process steps, and I shall continue to be mystified. Show me your process information model, and I won’t usually need your process steps.*”

3.1 Adoption Challenges and Information Model

Adopting a DevSecOps process in practice in an R&D organization is not straightforward. Table 1 summarizes the challenges faced at VAS in the adoption of the DevSecOps process, and how each challenge relates to the information model.

Table 1. Challenges in DevSecOps adoption

Challenge in DevSecOps process adoption	Challenge as related to the information model of the process
1. Poor developer and manager experience since relevant instructions are hard to find, understand, and follow, if only traditional, generic process descriptions are published	Traditional process descriptions are monolithic and long Lack of arrangements such as direct links that would direct the person to the relevant process instructions Process descriptions are static
2. Poor visibility to the current state of the DevSecOps and its adoption across the organization	Lack of adoption status in the process information model Traditional process information model cannot be queried
3. The maturity model of IEC 62443–4-1 suggests adopting simultaneously all maturity levels within the standard	Traditional process descriptions are not modular enough to support gradual adoption
4. It is hard to gather evidence of compliance for certification	Process information model usually does not directly accumulate evidence of practicing the process

3.2 Design Principle 1: Information Model Before Process/task View

Table 2 formulates the basic realization of the proposed solution/method into a design principle of clarifying an information model for the process before detailing the process tasks or steps. In the context of VAS' DevSecOps process, we instantiated design principle 1 with an information model that we call an issue graph. The artifacts of the process are called issues, which are linked together to form a graph. The issues may include, for example, process descriptions, project documentation, security-related issues, and internal controls to the content of the DevSecOps process.

Table 2. Design principle 1

Design principle	Information model before process/task view
Aim, implementer, and user	DevSecOps process practitioner: Facilitate process adoption and standard compliance of an R&D organization
Context	When an R&D organization aims to implement a DevSecOps process in compliance to a standard
Mechanism and enactor	Represent and communicate all the information related to the process with unified, modular, and interlinked information model tailored for the organization. Information model should be available throughout the organization including information artifact maintainers
Rationale	To answer the challenge #1 of Table 1, members of R&D organization need to understand all the information artifacts and their relationships to accept and understand the rationale and the descriptions of the tasks/steps in the DevSecOps processes

Figure 1 presents the issue graph information model using the entity relationship diagram (based on the notation by [3]). Issues are identified uniquely due to the traceability requirements of the standard, and we maintain a modification audit trail. The issues of the same type follow the same workflow state machine, and each issue is in a specific workflow state.

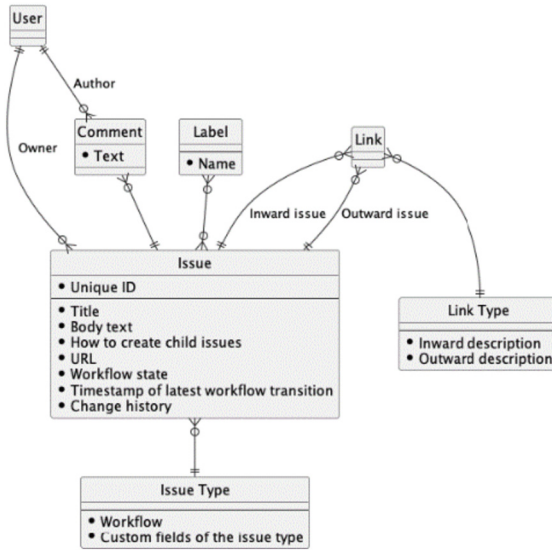


Fig. 1. An Entity Relationship Diagram of the “issue graph” process information model.

The issues of the graph are generated by creating new branches to the graph from template branches. Each issue may include instructions or means for the user to instantiate new issues as children of the said issue. Each issue is owned by a user, who is responsible for maintaining this piece of content. Users can collaborate on the issues by commenting on them and by referring to them by a URL. Issues can be tagged or labelled to categorize them for different metrics and to facilitate the process. The issues can be linked together with different types of links.

Each of issue types has an issue-type-specific workflow state machine, which may have issue-type-specific custom fields. The issue types of the VAS [IEC 62443–4-1] process model include controlled documents for process descriptions and project documentation, security issues that need to be managed, internal controls, security requirements, tests cases and test executions (Table 3).

Table 3. Issue types of the [IEC 62443–4-1] process information model

Issue type	Purpose	Custom fields
Controlled document	Describe a process description or a specification	–
User interface view	Shows instructions, metrics, and links in a specific context	–
Security issue	A security finding that needs to be managed	Affected component(s) and version(s), Fix version(s), Description of mitigation, Original risk level, Release urgency, Root cause category, Root cause analysis, Whether the issue should be disclosed to users, what kind of testing of the resolution is applicable
Internal control	A standard task of the DevSecOps process	Description of how to decide whether the control is in scope, Description of how to review whether control is OK, Last reviewed (timestamp)
Security requirement	A security-related requirement that has been recorded for a project	–
Test case	The description of how to test a security issue has been fixed or a security requirement has been implemented	–
Test execution	The description of the execution of a test case for a specific product version in a specific environment	–

The main link type is a descendancy relationship or parent-child relationship which produces a tree-based structure. The descendancy hierarchy can be used for access control, by giving different users read or write access to different branches of the tree. In addition to the parent-child links, there are other types of links that capture the relationship between issues (e.g., a security requirement may mitigate a security issue). A test case may be designed to verify a security requirement or mitigation of a security issue. The types of most important issue links are described in Table 4.

Table 4. Types of issue links

Link type	Purpose	Inward description	Outward description
Descendance	Links child issues to parents for the main tree hierarchy	Is parent of	Is child of
Mitigation	Links security requirements or implementation tasks to security threats	Is mitigated by	Mitigates
Template	Links issues to the templates from which they were created	Is instantiated as	Is created from template
Verification	Links tests cases to security requirements or security issues	Is tested by	Tests
Execution	Links test executions to test cases	Is executed by	Executes

Common feedback received from developers and managers in the retrospectives over the years was that it is hard to understand what should be done in practice and concretely to follow the standard compliant DevSecOps process. At the same time, we found that it is not feasible to provide very specific step-by-step instructions, because they would be too long and tedious to use and maintain. To our experience, organizing the information related to the DevSecOps process more clearly has helped developers and managers to get an overview of the process and understand what needs to be done.

3.3 Design Principle 2: Information Model Modularity

Many of the benefits of the information model are based on its modularity, which we have described as a separate design principle in Table 5. When design principle 2 was instantiated at VAS, we designed the issue types so that each issue type has a workflow state according to an issue type-specific workflow state machine. This makes the state of the information model searchable and enables the creation of various metrics and statistics.

Table 5. Design principle 2

Design principle	Information model modularity
Aim, implementer, and user	DevSecOps process practitioner: Facilitate <i>gradual, measurable, and sustainable</i> adoption of the DevSecOps process and standard compliance of an R&D organization
Context	When the R&D organization aims to implement a DevSecOps process in compliance to a standard
Mechanism and enactors	Modularity of the process information model, gradual small-step deployment of information model, and tracking the state of each artifact by the members of an organization
Rationale	Modularity makes gradual adoption easier, because individually adoptable items can be separated and from team's perspective, unnecessary views can be filtered. This addresses challenges #1 and #3 in Table 1 Measuring the organization's adoption state with simple measures from information model eases to manage the adoption of the process. This addresses the challenge #2 in Table 1 Simple and independent artifacts are easy to understand and manage by the enactors and they promote good management and development practices

We wrote small snippets of instructions, which we reused and included in both process descriptions and to the relevant contexts in various user interface views. This makes it easier to apply instructions that are relevant for the user in their current task. We also used a special issue type, internal control, which represents the standard DevSecOps tasks in a project. The internal controls can be labelled into separate adoption steps, so that a team can concentrate on a subset of the internal controls at a time. This helps with gradual adoption of the process.

The internal control is an issue type that models the standard tasks of the DevSecOps process. This concept is not included in the [IEC 62443-4-1] standard, and it is not used in all organizations that have a DevSecOps process. However, there are several benefits using internal controls: they help with gradual adoption of the process; they help with making the scoping decisions about which tasks are applicable and they enable a standard progress metric about the process adoption.

The workflow state machine of the internal control is shown in Fig. 2. The default state in the beginning is **Open**, which means that there is no decision whether the task is in scope for the project. Acceptable states are **Not Required** (task out of scope) and **OK** (task done). There is no end state, because DevSecOps is a continuous process. By tracking the last reviewed timestamp, controls in the **OK/Not Required** state can be highlighted to require attention. There is a state transition from the **OK** state to the same state, so that the last reviewed timestamp can be easily updated.

The notion of internal control was introduced in response to an R&D director's request, at the end of the first BIE cycle, to make the adoption of the DevSecOps directly measurable with simple metrics. We have also tracked the adoption of the DevSecOps

process by annual targets that we set based on metrics that we derived directly from the information artifacts: from the status of internal controls and security issues. To our experience, it is easier to lead the adoption, when R&D leaders can set measurable targets. Having a granular information model enables us to adopt the large standard compliant DevSecOps process gradually and easily at team level.

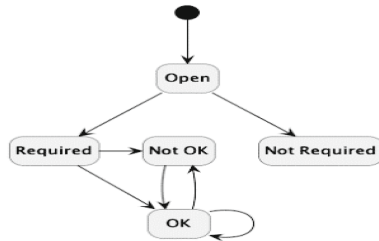


Fig. 2. Workflow state machine for internal control

3.4 Design Principle 3: Information Model Tailoring

In VAS, the DevSecOps process and thereby its information model needed to be tailored for the specific demands of the organization and in some cases even to the individual teams. The design principle of information model tailoring to the organization is presented in Table 6.

Table 6. Design principle 3

Design principle	Information model tailoring to the organization
Aim, implementer, and user	DevSecOps process practitioner: Facilitate adoption of the DevSecOps process and standard compliance of an R&D organization with as little friction as possible
Context	When the R&D organization aims to implement a DevSecOps process in compliance to a standard
Mechanism and enactors	The process information model (incl. The content) should be tailored/mapped based on organization’s existing tools and practices. The practical process and tooling used must maintain the integrity of the data model automatically or manually The information model may require team-wise tailoring where a reference information model can be used as a starting point
Rationale	The mappings of the information model to concrete development tools act as integration points to integrate DevSecOps practices with tools/practices that development team is using. By maximizing the use of existing tools/practices, changes within the organization can be minimized that addresses the challenge #1 in Table 1. Mapping practices to tools ensures accumulating the process to the tools when enactors apply the process. This addresses the challenge #4 in Table 1

The VAS’ DevSecOps process implements its issue graph information model mainly based on the Atlassian Confluence and Jira tools that have been an inspiration for many characteristics of the information model. These tools have limitations so that not all steps

described here could be automated and, thus, have influenced the design of the hierarchy of the issue graph. Table 7 illustrates the Valmet Automation implementation.

Table 7. Implementation of the issue types of the issue graph model in VAS DevSecOps

Issue type	Implementation at VAS
Controlled document	Confluence pages and templates
User interface view	Confluence pages
Security issue	A custom Jira/Azure DevOps backlog issue type
Internal control	A custom Jira issue type called SDL control
Security requirement	A custom Jira issue type and a Jira issue template that include four review subtasks
Test Cases and Executions	Jira Xray test cases/executions for manual/semi-automatic tests. Robot framework test cases/executions for automated tests

When we added new VAS teams to adopt the DevSecOps process after the first BIE cycle, we noticed immediately that the information model must be tailored team-wise and to fit the needs of projects of different sizes and different technology scopes.

As part of the Insta customer interviews during the fourth BIE cycle, we learned that the participants of the interviews preferred integrating security practices to their existing tools. The challenges of easily finding evidence of practicing a standard-compliant process are obvious to anyone who has had their process audited for certification. To our experience, it pays off to configure the tools that developers and managers already use so that evidence is accumulated automatically to the tools.

3.5 Design Objectives

Not all teams use the Atlassian tools for managing their work. Many use, for example, Azure DevOps, and Office for documentation. This justifies the design principle of information model tailoring, as a root cause for a lot of tedious work for DevSecOps process practitioners who try to support these teams. Documents and document templates must be converted between tool specific formats, and similar information needs to be maintained in multiple places. A portable representation format for the information model could help with these challenges, as a design objective for the future (Table 8). Our other design objective (Table 9) proposes to develop a process information model tool that would help with keeping the information model coherent across different tools. While there are existing tools for managing backlogs and tracking issues, tools for development documentation in an enterprise wiki, and tools for modeling the software architecture, the authors are not aware of any existing tools for managing the information model for a DevSecOps process.

Table 8. Design objective of portable information model representation

Design objective	Process information model representation
Aim, implementer, and user	DevSecOps process practitioner: Maintain the DevSecOps process or facilitate sharing of best practices for process adoption between organizations/teams that use different tools
Context	When the R&D organization aims to implement a DevSecOps process in a compliant to a standard
Mechanism and enactors	A portable process information model as a tool-independent presentation. The representation based on common file formats (e.g., YAML and JSON), stored in version control and processed by scripts. DevSecOps process practitioner maintains a reference process information model in the portable format and generates tool-specific representations by (semi)-automated means
Rationale	A portable representation used as a single source, from which other versions can be derived. Resource savings by maintaining the process information models for different teams. Reduced errors aligned information models. Portable representations stored in a software version control system, which supports audit trail and enables open collaboration using similar techniques as open-source software projects use, such as change request reviews. Several organizations can collaborate on the process information model. In the future, a clearly defined textual representation format lends itself to the use of generative artificial intelligence, for example for providing suggestions to security specifications

Table 9. Design objective of process information model management tool

Design objective	Process information model management tool
Aim, implementer, and user	DevSecOps process practitioner: Facilitate the process adoption
Context	When the R&D organization aims to implement a DevSecOps process in compliance to a standard
Mechanism and enactors	A process information model management tool is developed for maintaining the consistency of the process information model in other tools. The tool could utilize the model described in Table 8. It could be used for instantiating the process information model or new branches of the information model and keeping the model consistent across tools. The commonly used R&D issue trackers and documentation tools have an API, which the process information management tool could use to create and maintain the issues and documents
Rationale	Automating the repeating tasks in process information model management improves user experience, reduces workload, and decreases the possibilities for errors

4 Conclusions

The challenges of adopting DevSecOps and security standard compliance in agile OT development have been acknowledged in recent academic literature by Akbar et al. [1] and Rajapakse et al. [14] as well as in industry-oriented articles by Moyon et al. [10–12]. In this paper, we outlined one of the first experience reports on adopting and implementing IEC 62443 standard in practice. Based on the experiences, we proposed three design principles for standard compliant DevSecOps practices for OT development. These design principles originate from observations and experiences of several years in cybersecurity and software development practice in OT industry. The main influence on the creation of these design principles is that the same challenges or problems are encountered in numerous companies with minor variations. Altogether, the experiences suggest for an information-centric view on adopting and using the 62443–4–1 standard with DevSecOps to precede and complement the previously suggested process/task-centric view by Moyon et al. [10–12]. The information-centric view suggests that a shared information model of security issues gives common ground while allowing for more contextual, actual processes to integrate security work in DevOps. Such a common information model enables sharing, coordination, and reporting of security issues even when DevSecOps is implemented through often varying tasks across team-specific development processes and tools.

The information model behind the design principles emerged through the ADR BIE cycles to provide a theoretical background for the proposed solution. Design principles set up general guidelines on the adoption and implementation of IEC 62443, accelerating the operationalization of the standard into practice. As a case study, we described how the design principles are instantiated at VAS, while the formulation of the design principles suggests for their applicability beyond the case study at hand. Besides the design principles, we proposed two design objectives for the future: portable process information model representation and process information model management tool. These objectives are needed to address such technical questions as conversion between different formats and coherence of information model across tools.

This information-centric approach to adopt and implement complex standards such as IEC 62443 into practice complements the previously proposed process/activity-centric approach. Solutions in this paper are constructed abductively from empirical observations and development experiences to theory direction. The proposed approach sets up a new direction to the adoption and implementation of the requirements of IEC 62443 into practice and fulfils the hitherto addressed gap of missing experience reports in the scientific literature.

Acknowledgements. To Janne Ahlberg, Suvi Kaartinen, Joonas Lepola, Elina Niemimaa, Kim Paananen, Mikko Tervo and Markku Tyynelä for contributions to the DevSecOps framework. This study was funded in part by Business Finland (ITEA3 project OXILATE, <https://itea4.org/project/oxilate.html>).

Disclosure of Interests. Authors 1 and 4 were employed by Insta and author 3 by VAS during the research. Author 2 has no competing interests in the contents or results of this research.

References

1. Akbar, M.A., Smolander, K., Mahmood, S., Alsanad, A.: Toward successful DevSecOps in software development organizations: a decision-making framework. *Inf. Softw. Technol.* **147**, 106894 (2022)
2. Brooks, F.P., Jr.: *The Mythical Man-Month*. Addison-Wesley, Reading MA (1975)
3. Everest, G.C.: *Database Management: Objectives, System Functions, and Administration*. McGraw-Hill, New York (1986)
4. Givvehchi, O., Landsdorf, K., Simoens, P., Colombo, A.W.: Interoperability for industrial cyber-physical systems: an approach for legacy systems. *IEEE Trans. Industr. Inf.* **13**(6), 3370–3378 (2017)
5. Gregor, S., Chandra Kruse, L., Seidel, S.: Research perspectives: the anatomy of a design principle. *J. Assoc. Inf. Syst.* **21**(6), 1622–1652 (2020)
6. Iden, J., Tessem, B., Päiväranta, T.: Problems in the interplay of development and IT operations in system development projects: a delphi study of Norwegian IT experts. *Inf. Softw. Technol.* **53**(4), 394–406 (2011)
7. IEC 62443-4-1 International Standard: Security for industrial automation and control systems - Part 4-1: Secure product development lifecycle requirements. International Electrotechnical Commission (2018)
8. ISASecure: ISASecure SDLA Certified Development Organizations. International Society of Automation. <https://isasecure.org/end-users/iec-62443-4-1-certified-development-organizations>. Accessed 23 Dec 2022
9. Leite, L., Rocha, C., Kon, F., Milojicic, D., Meirelles, P.: A survey of DevOps concepts and challenges. *ACM Comput. Surv.* **52**(6), 1–35 (2020). <https://doi.org/10.1145/3359981>
10. Moyón, F., Almeida, P., Riofrío, D., Mendez, D., Kalinowski, M.: Security compliance in agile software development: a systematic mapping study. In: *Proceedings of the 2020 46th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pp. 413–420 (2020)
11. Moyón, F., Soares, R., Pinto-Albuquerque, M., Mendez, D., Beckers, K.: Integration of security standards in DevOps pipelines: an industry case study. In: *Morisio, M., Torchiano, M., Jedlitschka, A. (eds.) Product-Focused Software Process Improvement (PROFES 2020)*. Lecture Notes in Computer Science, vol. 12562, pp. 434–452. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64148-1_27
12. Moyón, F., Méndez, D., Beckers, K., Klepper, S.: Using process models to understand security standards. In: *Bureš, T., Dondi, R., Gamper, J., Guerrini, G., Jurdziński, T., Pahl, C., Sikora, F., Wong, P.W.H. (eds.) SOFSEM 2021*. LNCS, vol. 12607, pp. 458–471. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67731-2_34
13. Myrbakken, H., Colomo-Palacios, R.: DevSecOps: a multivocal literature review. In: *Mas, A., Mesquida, A., O'Connor, R., Rout, T., Dorling, A. (eds.) Software Process Improvement and Capability Determination (SPICE 2017)*. Communication in Computer and Information Science, vol. 770, pp. 17–20. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67383-7_2
14. Rajapakse, R.N., Zahedi, M., Babar, M.A., Shen, H.: Challenges and solutions when adopting DevSecOps: a systematic review. *Inf. Softw. Technol.* **141**, 106700 (2022)
15. Rindell, K., Ruohonen, J., Holvitie, J., Hyrynsalmi, S., Leppänen, V.: Security in agile software development: a practitioner survey. *Inf. Softw. Technol.* **131**, 106488 (2021)
16. Sein, M.K., Henfridsson, O., Purao, S., Rossi, M., Lindgren, R.: Action design research. *MIS Q.* **35**(1), 37–56 (2011)



Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Exploring Emotions in Online Team Meetings: Unpacking Agile Retrospective

Dron Khanna¹(✉)  and Abdullah Aldaej² 

¹ Free University of Bozen-Bolzano, Bolzano 39100, Italy
dron.khanna@unibz.it

² Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
aaaldaeej@iau.edu.sa

Abstract. Establishing a psychologically safe work environment is crucial for leading a positive and practical agile retrospective. Emotions are closely intertwined concepts that come under the roof of psychology. Capturing them at the right time helps to detect harmful or favourable online behaviours, hinder or facilitate the software development cycle, and moralize or demoralize the team in a software company. This study aims to identify emotions that appear during the online agile retrospective. Our study asks the research question: How often are different emotions repeated during the online agile retrospective? We conducted a multiple case study with two software companies. We analyzed three recorded online retrospective sessions to seize various emotions. Our findings show that eighteen emotions appear on the agile retrospective. Some of the highest repeated emotions are approval, realization, excitement, relief, disappointment, confusion, optimism, and disapproval.

Keywords: Emotions · Agile retrospectives · Online meetings · Online Teams · Retrospectives

1 Introduction

The software development landscape continuously evolves, and agile methodology has delegated teams to adapt and deliver value in the dynamic work environment. Agile retrospectives are a capstone of the agile framework and a crucial practice to many software development teams [1]. The human element must be noticed in the agile retrospective cycle as it directly affects the success of the software development cycle [2]. Establishing a psychologically safe work environment is crucial for leading to positive and practical agile retrospective sessions. When the team reflects on the experience, areas needing improvement, and action plans at the end of each iteration, they express themselves by sharing thoughts and opinions [3]. While doing the same, psychological safety elements, i.e. emotions, are involved during the online retrospective meeting [4]. A couple of words expressed during the online meeting could lead to a negative or positive work environment [5]. Capturing emotions could be fruitful as it helps to detect harmful or favourable online behaviours [6], hinder or facilitate the software

development cycle, moralize or demoralize the team, and nourish or discourage innovation and cooperation inside the organization. Hence, it is essential to gather the emotions of the agile retrospective teams [4]. Emotions are interconnected and related concepts placed under the roof of human psychology and communication. Emotions can be a range of feelings, for example, *happy, sad, anger, fear, etc.* [11]. The structure of emotions is the basis for creating human sentiment [7]. So, sentiment is considered the high-level category of emotions, categorized into three categories: positive, neutral, and negative. Over time, emotions are connected with certain experiences and beliefs that generate a sentiment [8]. This study revolves around emotions during the agile retrospectives of two software development teams. We aim to investigate the various emotions contributing to agile teams. Our research question is: **Rq.) How often are different emotions repeated during the online agile retrospective?** We conducted multiple case studies to detect the type of emotions and their frequency in the retrospectives from two software development teams. We found that several emotions, such as (*approval, realization, excitement, relief, disappointment, etc.*) overlap in both the agile retrospectives. *Approval* was repeated maximum (17 times) whereas *pride, fear, embarrassment* was minimum as it occurred only once.

2 Background and Related Work

2.1 Emotions in Online Agile Retrospectives

Software teams at the workplace express many emotions that impact their productivity. Girardi et al. investigate the correlation between developers' emotions and productivity. The authors experimented with 21 developers from five Dutch software companies [9]. The study identified a positive correlation between developers' emotions and perceived productivity. In addition, Graziotin et al. examine the effect of emotions experienced by software developers [10]. Based on the survey results from 317 participants, the authors found that emotions have some impact related to the happiness and unhappiness of developers. These developers practising retrospectives should feel psychologically safe [3], which encourages them to share their experiences and emotions [4]. A recent study by Grassi et al. describes the importance of emotions in agile retrospectives and how students' emotions vary through performing activities in a software engineering course. The authors developed an emotion visualization tool that visualizes emotions, actions, and bio-metrics. Agile retrospectives were chosen as a test bed to evaluate the tool. The study shows that detecting emotions can assist in discussing and fixing various issues that arise in a sprint [4]. However, there needs to be more research that applies emotion analysis in online agile retrospective meetings. Often, it is noticed that retrospective participants use emojis to express emotions at various stages of the meeting, for example, during a chat [3].



Fig. 1. Theoretical Framework

2.2 Theoretical Framework

This section specifies various emotions that we collected from literature [11–14], which serve as the base of the theoretical framework. **Neutral emotions:** Neutral emotional states that are neither positive nor negative. The literature outlines three **neutral emotions** and their following example. 1.) **Confusion:** “Ok, just making sure I was confused”, 2.) **Curiosity:** “I am curious to know about [something]”, 3.) **Realization:** “I figured/realized something” [11,12]. **Positive emotions:** These emotional states or reactions are welcoming, nice, inspiring, and delightful. The literature outlines eleven **positive emotions** and their following example. 1.) **Admiration:** “Please keep up the great work”, 2.) **Amusement:** “Haha, actually, grandpa did! Go figure”, 3.) **Approval:** “We have received approval from the boss”, 4.) **Caring:** “He was caring for his dog”, 5.) **Desire:** “I can’t wait to hear the stories”, 6.) **Excitement (Gratitude, Joy, Enthusiasm):** “Excellent idea, thank you”, 7.) **Love (Affection, Adoration, Cuteness):** “Cause you were so tiny and fragile”, 8.) **Optimism:** “I am confident about it”, 9.) **Pride:** “We are the best”, 10.) **Relief:** “Thank god, I was just thinking to do it”, 11.) **Surprise:** “Wow, what a sunny day” [13,14]. **Negative emotions:** Negative emotions are reactions that are unwelcoming, unpleasant, upsetting, and uneasy. The literature outlines eleven **negative emotions** and their following example. 1.) **Anger:** “If this is who you are”, 2.) **Annoyance:** “But the man keeps it tearing apart”, 3.) **Disapproval:** “She is not ready yet”, 4.) **Disappointment:** “vmware fusion seems to get slower and slower”, 5.) **Disgust:** “Well that made me want to continue to live in Alberta”, 6.) **Embarrassment:** “I feel foolish”, 7.) **Fear (Anxiety, Nervousness):** “Is Someone there”, 8.) **Grief (Pain, Tiredness):** “It’s back, what I mean is my headache”, 9.) **Remorse (Guilt):** “I am sorry, I wasn’t perfect”, 10.) **Sadness (Distress):** “Poor guy”, 11.) **Surprise:** “What, you won’t be two blocks away anymore?” [11,14]. As shown in Fig. 1, we captured similar examples of emotions and mapped them with the (audio, text and icon) involved in agile retrospective meetings. With the help of the framework, we retrieved a list of emotions and their frequency presented in the agile retrospective.

3 The Study Research Process

We conducted multiple case studies to collect emotions from two software development teams. We selected the cases based on the convenience sampling approach [17]. The first case is a team (T1), a software company based in Germany

that helps to calculate assessment management for many individuals and organizations. The second team (T2) is a multinational software company with several European offices. **Data collection** - We collected the data from three online retrospective meeting videos. (M). One video from (T1) and two videos from (T2). The team (T1) meeting (T1-M1: lasted around 35 min with 3 participants and used [Trello](#) and Zoom as software tools for OAR). From the two videos of T2, we used the first one as our pilot study (T2-M1: lasted around 15 min with 10 participants) and the other one as our case (T2-M2: lasted approximately 30 min with 10 participants and used a digital board called [Parabol](#), Microsoft Teams). We converted the videos into text through [cockatoo](#) (software that converts video to text files). The text was used as our transcripts for analysis.

Table 1. Data Analysis

Chunk	Time-stamp	Neutral Emotion	Positive Emotion	Negative Emotion
1	0.04 0.12 0.48	Curiosity	Happy, Excitement, Desire Approval Amusement	
2	1.32			Disappointed
3	3.1	Realisation		

Data Analysis. We applied the research approach called the “bracketing technique” to analyse the three videos [15]. This technique helps to describe precise time-stamped breakpoints and use them for coding. First, we analysed the pilot study (T2-M1), and later, we completed the analysis of T1-M1 and T2-M2 meetings. To analyse each time-stamp or chunk (1 min long), the authors manually listened to the audio first and then validated the text with the theoretical framework. Both authors together picked each minute chunk (few examples are visible in Table 1) one by one (chunks 1,2, and so on), assigned emotion labels based on the theoretical framework, and reached a consensus on the identified emotion. Although the retrospectives lasted for around 30–35 minutes, we found only 28 min of instances or chunks for T1 and 17 chunks for T2 due to the following reasons. We excluded chunks were: 1.) Teams had no audio content relevant to the retrospective available that could be converted to text; 2.) The team reflected or thought during the period; hence, no conversation or text was shared during the meeting. After analysing the manually collected emotions of the chunks, we first used software tool [Text2data](#) and then [ChatGPT-3.5](#) to analyse the text and validate our results. We found out that our study was similar to ChatGPT analysis compared to the tool. We discovered that the tool only used minimal emotions to calculate compared to our manual calculations based on theoretical framework. The software tool considered only fifteen types of emotions: “anger, boredom, emptiness, enthusiasm, fear, fun, happiness, hate, joy, love, neutral, relief, sadness, surprise, and worry”. In contrast, the theoretical framework 2.2 in the previous section consists of twenty-five types of emotions. We also asked ChatGPT what methods and algorithms generated the

analysis. The ChatGPT used the following (satisfaction, happiness, agreement, contentment, joy, approval, optimism, dissatisfaction, concerns, criticisms, disappointment, frustration, scepticism) emotions for the analysis.

4 Findings

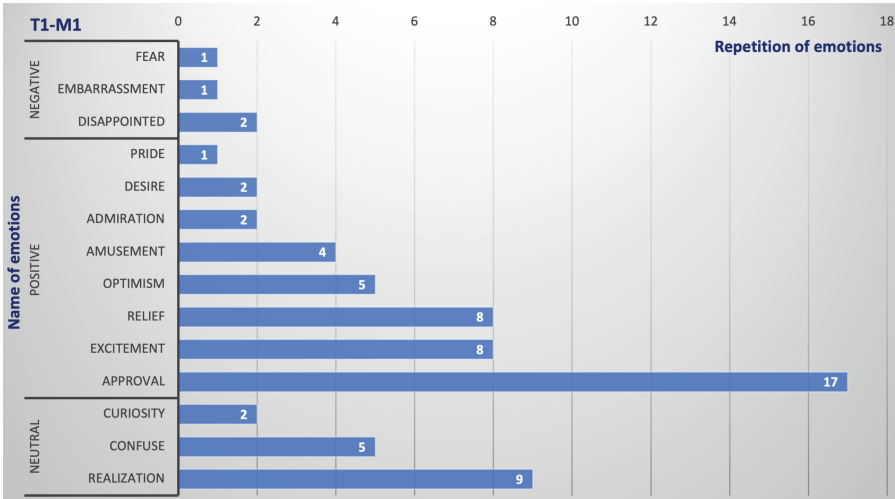


Fig. 2. Name of the emotions and their repetition in T1-M1 retrospective

Figures 2 and 3 present the type of emotions in the online retrospective meetings. In both figures, the X-axis represents the number of (times) or frequencies the emotions were repeated, and the Y-axis represents the type (name) of the emotion.

Neutral emotions: We can observe that **Realization (9 times repeated) and Curiosity (2 times repeated)** were the two common neutral emotions in retrospectives. It shows that retrospective members were either realizing or curious about the sprint’s past, present, or future tasks. For example, a participant realized: *“we didn’t probably think it through completely. We ended up completing it. But probably in the other direction, so now we have to consider it and the next steps to see how we can go back on our steps. Let’s go on with one of the other cards.” (T1-M1)*. Whereas another team member was curious: *“OK. Can we put the [task] inside the sprint?”, “So maybe we can start with the collaboration and coordination between the two teams if you agree?”* and realized: *“OK, got it, so probably we should discuss the two deltas and this one here if you want.” (T2-M2)*.

Positive emotions: Observing the Figs. 2 and 3 positive emotions, **Approval (17 times repeated)** was the most preferred whereas the second

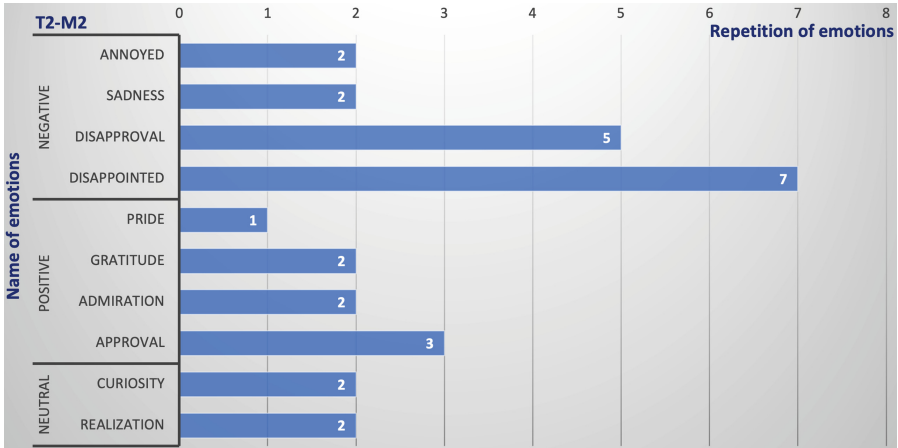


Fig. 3. Name of the emotions and their repetition in T2-M2 retrospective

most was **Excitement and Relief** from T1 and **Admiration** from T2 retrospective. Regarding **Approval**, one participant mentioned: “*Yeah, totally agree with that. I mean, we are both doing.*” “*We could be good at this one for an action point, right? What do you say? Yeah. What could we do this, actually.*” (T1-M1). The same team was also excited and relieved: “*I’m delighted. Three people have already said yes. So, I’m pretty happy with it. Yeah. And this can lead to a lot better estimates.*” (T1-M1). The second team encountered an admiration moment where the participant quoted “*Thank you to [Names] for your continuous patience and help during this sprint. [Names], best teammates ever, and thanks to have followed the DB activity.*” (T2-M2).

Negative emotions: Concerning the negative emotions, both teams had a **Disappointed (on team 7 times repeated)** feeling. Second repeated, **Embarrassment or Fear** for T1 and **Disapproval** for T2 as a negative emotion during the retrospective. The team was **disappointed** and quoted “*Delta, team A, and B working on the same project, with no coordination at all. Delta, Are eight story points issues too big? How can we avoid the failure of the sprint?*” (T2-M2). The team T1 had a **fear** about estimation as they mentioned “*So, let’s just be careful. Yeah, it is affecting too much the planning for Q1 for those estimations?*” (T1-M1) Whereas there was a moment of **disapproval** as one member mentioned “*No, I disagree with this. So this was not the idea of the teams, I think. No. The teams should be independent.*” (T2-M2).

5 Discussion

Our study sheds light on various emotions during the online agile retrospective. Emotions are intrinsic to human communications, and our findings suggest that they can help retrospective groups shape better outcomes and learnings.

Within the two software case study teams, we found in total eighteen emotions (see Figs. 2 and 3), namely [3 neutral (*Realization, Confuse, Curiosity*), nine positive (*Approval, Excitement, Relief, Optimism, Amusements, Admiration, Desire, Pride, Gratitude*), and six negative (*Disappointed, Disapproval, Sadness, Annoyed, Embarrassment, Fear*)] emotions. We also identified the overlap of various emotions between the two cases such as (*Realization, Curiosity, Approval, Admiration, Pride, and Disappointed*). Knowing the emotions can help to encourage psychological safety, strengthen empathy, and generate pain points and insight in a team. But to grasp the emotions, the team must respect confidentiality and treat all the members with respect in the company. We observed that factors like the company's culture and the scrum leader's behaviour facilitating the retrospective could influence emotions. Moreover, a tone could also affect the comfort level of participants and change of mindset to discuss the task positively or negatively. This study lays direct implications for agile practitioners. Retrospective teams can create an environment to encourage communications with open expression of positive emotions and constructively managing negative emotions. Teams could focus better on the improvements of a cycle and apply some methods to solve the negatively evoked issues before the end of the retrospective. A team could use tools, as mentioned in the study [4], that could capture emotions during retrospective sessions. Concerning the limitation of this study. It was conducted with only three retrospective videos. We had a limited number of videos because retrospectives are a practice that occurs at the end of the sprint cycle [16], but usually, it is longer than other meetings. Hence, we selected an agile retrospective for the study. This limits the generalizability of our findings. Future research could involve additional sessions of retrospectives, sprint planning, daily planning, daily stand-up, and product feedback that could lead to a better understanding of both sentiments and emotions in online agile retrospectives.

6 Conclusion

Human emotions are the factors that affect the success of agile retrospectives. In this paper, we study the emotions in online agile retrospectives from two software teams by identifying how often emotions are repeated throughout the agile retrospective. Our study reveals that approval, excitement, admiration, and relief are the most positive emotions. Disappointment and Disapproval are the most frequent negative emotions. At the same time, realisation and curiosity account for neutral emotions. Emotions are crucial in shaping the digital interaction, team dynamics and decision-making process. Revealed emotions act as a facilitator that affects the performance of a team. It is vital to foster the trend of psychological safety in agile retrospectives so that teams in organizations can boldly express their emotions, leading to improved sprint cycles. In the future, the additional research should encompass sentiments obtained from emotions, which could further enhance the entire software development process.

References

1. Turetken, O., Stojanov, I., Trienekens, J.J.: Assessing the adoption level of scaled agile development: a maturity model for scaled agile framework. *J. Softw. Evol. Process* **29**(6), e1796 (2017)
2. Derby, E., Larsen, D., Schwaber, K.: *Agile Retrospectives: Making Good Teams Great*. Pragmatic Bookshelf (2006)
3. Khanna, D., Wang, X.: Are your online agile retrospectives psychologically safe? The usage of online tools. In: Stray, V., Stol, K.J., Paasivaara, M., Kruchten, P. (eds.) *Agile Processes in Software Engineering and Extreme Programming. XP 2022. Lecture Notes in Business Information Processing*, vol. 445, pp. 35–51. Springer, Cham (2022)
4. Grassi, D., Lanubile, F., Novielli, N., Serebrenik, A.: Towards supporting emotion awareness in retrospective meetings. In: *2023 IEEE/ACM 45th International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*, pp. 101–105. IEEE (2023)
5. Zhou, K., Constantinides, M., Joglekar, S., Quercia, D.: Predicting meeting success with nuanced emotions. *IEEE Pervasive Comput.* **21**(2), 51–59 (2022)
6. Ayoko, O.B., Konrad, A.M., Boyle, M.V.: Online work: managing conflict and emotions for performance in virtual teams. *Eur. Manag. J.* **30**(2), 156–174 (2012)
7. Frijda, N.H., Kuipers, P., Ter Schure, E.: Relations among emotion, appraisal, and emotional action readiness. *J. Pers. Soc. Psychol.* **57**(2), 212 (1989)
8. Bagozzi, R.P., Pieters, R.: Goal-directed emotions. *Cogn. Emot.* **12**(1), 1–26 (1998)
9. Girardi, D., Lanubile, F., Novielli, N., Serebrenik, A.: Emotions and perceived productivity of software developers at the workplace. *IEEE Trans. Software Eng.* **48**(9), 3326–3341 (2022)
10. Graziotin, D., Fagerholm, F., Wang, X., Abrahamsson, P.: What happens when software developers are (un)happy. *J. Syst. Software* **140**, 32–47 (2018)
11. Cowen, A.S., Elfenbein, H.A., Laukka, P., Keltner, D.: Mapping 24 emotions conveyed by brief human vocalization. *Am. Psychol.* **74**(6), 698 (2019)
12. Cowen, A., Sauter, D., Tracy, J.L., Keltner, D.: Mapping the passions: toward a high-dimensional taxonomy of emotional experience and expression. *Psychol. Sci. Publ. Interest* **20**(1), 69–90 (2019)
13. Cowen, A.S., Keltner, D.: Semantic space theory: a computational approach to emotion. *Trends Cogn. Sci.* **25**(2), 124–136 (2021)
14. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S.: Goemotions: a dataset of fine-grained emotions. arXiv preprint [arXiv:2005.00547](https://arxiv.org/abs/2005.00547) (2020)
15. Gearing, R.E.: Bracketing in research: a typology. *Qual. Health Res.* **14**(10), 1429–1452 (2004)
16. Khanna, Dron, Nguyen-Duc, Anh, Wang, Xiaofeng: From MVPs to pivots: a hypothesis-driven journey of two software startups. In: Wnuk, Krzysztof, Brinkemper, Sjaak (eds.) *ICSOB 2018. LNBIP*, vol. 336, pp. 172–186. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04840-2_12
17. Baltes, S., Ralph, P.: Sampling in software engineering research: a critical review and guidelines. *Empir. Softw. Eng.* **27**(4), 94 (2022)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Emerging Digital World



Feeling the Elephant: Insiders' Perspectives on the Metaverse

Fabrício de Oliveira¹ , Xiaofeng Wang² , and Luciana Zaina¹  

¹ Federal University of São Carlos, Sorocaba, São Paulo, Brazil
fabricio.malta@estudante.ufscar.br , lzaina@ufscar.br

² Free University of Bozen-Bolzano, Bolzano, Italy
xiaofeng.wang@unibz.it

Abstract. The metaverse has been considered in various literature reviews as a multifaceted and complex concept that can not be defined from a single set of terms. These literature reviews have attempted the Metaverse definition based on the research most published before the heated attention on the Metaverse in 2021; therefore, they may not provide an up-to-date understanding of the phenomenon that incorporates the perspectives from the industry. This paper aims to disentangle the complexity of the Metaverse concept considering the perspectives of insiders - practitioners who play essential roles in the recent Metaverse wave. To achieve our goal, we analyzed one specific type of gray literature - a podcast series from Bloomberg entitled “Into the Metaverse” which featured different professionals active in the Metaverse landscape. Three themes were identified that represent the essential characteristics of the Metaverse which include technology capabilities, infrastructure characteristics, and social and economic aspects. Our study contributes to a more contemporary industrial understanding of the Metaverse concept. The understanding can assist researchers in future investigations into the evolving Metaverse paradigm.

Keywords: Metaverse · Grey literature · Thematic analysis · Industry perspective

1 Introduction

The Metaverse landscape has rapidly evolved in recent years since Facebook’s announcement of its transformation into Meta. Google Trends data underscore the surge in Metaverse-related searches since late 2021, reflecting its growing significance [8]. In tandem with the burgeoning popularity of the Metaverse, there has been a commensurate increase in the volume of related publications during the same period [1, 3].

Scholars and practitioners alike grapple with the challenge of defining the evolving concept of the Metaverse [2, 11]. The Metaverse represents an extension of the internet’s evolution, with the potential to merge seamlessly with our physical world through technologies like virtual reality (VR) and augmented reality

(AR) [1]. As industries like retail and entertainment venture into the Metaverse, the need for a universally accepted definition becomes increasingly pressing to facilitate interdisciplinary discourse [10]. However, due to the multifaceted nature of the Metaverse, its precise definition remains elusive.

Defining the Metaverse is like feeling an elephant. Different perspectives from different stakeholders existed. Each company interprets the Metaverse according to its needs, goals, and industry sectors, resulting in various definitions and applications [11]. While some companies may view the Metaverse as a platform for social interactions and entertainment experiences, others consider it a fundamental tool for future¹. This variety of interpretations highlights the need for a deep and context-specific analysis of the various definitions of the Metaverse prevalent in the industry [16].

Taking into account the motivations above, our study aims to fill the existing gaps in the current literature by adding a contemporary industrial understanding of the Metaverse concept. To this end, we formulated the following research question: *RQ: What is the definition of the Metaverse from the perspectives of the involved practitioners?*

To answer this question, we qualitatively explored data collected from 10 podcast episodes focused on the Metaverse. Other studies have investigated the Metaverse complex definition from the scientific perspective as in [2, 4], or considering few amounts of papers which showed professionals' viewpoint about this definition [1] or then did not explore the practitioners' perspective deeply [2, 11]. Our work differs from the others by taking into account the perspectives of insiders - professionals who are working in the fields related to the Metaverse and actively shaping its development.

The contribution of this paper is to bring forward the current understanding of the Metaverse concept from these insiders, which both researchers and practitioners can then use to make better sense of the elephant - the complex phenomenon of the Metaverse. Our findings restate some topics that have been discussed by other authors. Additionally, we uncovered new perspectives of people expressing themselves from avatars, and also emergent themes of discussion such as technologies for connectivity and the problem of distinguishing the real and virtual world.

The remainder of the paper is organized as follows: Section 2 presents related work that aims to define the Metaverse. Next, Sect. 3 describes our research method, and Sect. 4 presents the study's results. In Sect. 5, we discuss the results in response to the research question. Finally, we conclude the work in Sect. 6.

2 Related Work

The metaverse complex definition has been explored mainly from the scientific literature. The literature reviews have presented the Metaverse conception from a more broad and general perspective [2, 7, 10], while others focus on specific

¹ <https://tech.facebook.com/reality-labs/2021/10/connect-2021-our-vision-for-the-metaverse/>.

domains such as education where Metaverse has been strongly adopted [3,4,6]. There is also work aspiring the unified definitions by adopting an ontology to explain the Metaverse concept [5]. Little works have explored sources that explore the Metaverse definition from the perspective of practitioners [1,11].

The systematic literature review conducted by Ritterbusch and Teichmann [7] led to the understanding of the Metaverse as a decentralized, three-dimensional online environment that is both persistent and immersive. In the authors' perspective, users who are embodied by avatars and can interact socially and economically in virtual spaces that exist independently of the physical world. Considering 30 papers in a literature review of Chen et al. [9] stated that the definition of the Metaverse is mainly divided into two categories: service-related to the Metaverse and technology used in the Metaverse. For the service-oriented, the authors found that in the Metaverse, the avatar that represents users, the daily communication, and the community are essential and also allow real-time social interactions for many users simultaneously. In the techniques-oriented category, the Metaverse is seen as the next generation of the Internet, building a 3D virtual world using technologies like AR, VR, and MR and exploring blockchain as an economic system with virtual money.

Similarly, Almoqbel et al. [2] conducted a systematic literature review and considered service and technology perspectives to define four categories that represent the main characteristics of the Metaverse. The categories include activities, content creation, users and their roles, and technical specifications. Space was an additional theme (i.e., out of the scope of the main categories) which represents the most challenging and inconsistent topic. It points out different perspectives on the relationship between the Metaverse and the real world. Park and Kim [10] proposed concepts, and techniques for realizing the Metaverse from the analysis of 260 papers. These concepts and techniques are divided into three components: hardware, software, and content. According to the authors, hardware is crucial for creating immersive experiences, with Head-Mounted Displays (HMDs) serving as key devices. Software components encompass functions related to recognition and rendering. Content covers multimodal content representation, avatar modeling, and scenario generation population and evaluation.

Education emerged as an eminent application field of the Metaverse. Zhang et al. [4] defined it as an enhanced environment that fuses Metaverse-related technologies with elements of both virtual and real educational settings. According to the authors, this environment allows learners to use wearable devices to access education from anywhere, interact with various digital elements, and feel as if they are present in a physical classroom. The authors propose a framework for the Metaverse in education that highlights key technological components like high-speed communication and networks and technologies for managing computing analytical, modeling interaction and authentication.

In another work, Hwang and Chien [6] stated the Metaverse as an encompassing virtual environment with numerous applications in education providing learners with immersive, entertaining, and continuous experiences. It includes an authentic world for working and learning alongside intelligent non-player char-

acters (NPCs), tutors, peers, tutees, and other human learners. For the authors, the Metaverse topic presents challenges related to technology, ethics, and pedagogy. [3] analyzed 19 papers published between 2009 to 2022 from a qualitative approach. The results showed that in the late 2000s to mid-2010s the Metaverse was described as 3D digital virtual worlds where individuals could live and build their identities through avatars. After the mid-2010s, the definition remained relatively similar; however, it also encouraged communication, interaction, and collaboration among the users. For the author, the Metaverse is continuously evolving with advancements in technologies like AR, VR, and AI applied in learning environments. The author also proposed key elements to enhance the value of Metaverse for educational purposes that include immersion, advanced computing, socialization, and decentralization.

Abu-salih [5] employed the Design Science Research Methodology (DSRM) to design a domain ontology (MetaOntology) for the Metaverse. The resulting definition of the Metaverse is a digital ecosystem that encompasses advanced technologies and infrastructure. This ecosystem includes digitization aspects, key technologies (e.g., Virtual Reality, Augmented Reality), software and hardware components, metaverse content, tech companies, physical counterparts, and user feedback.

Different from the previously discussed work, Weinberger [1] included two non-academic publications in their work. The author conducted a meta-synthesis of both scientific literature and grey literature to provide a single Metaverse definition. This unified definition covers the themes of ubiquitous space, virtual worlds, use of avatars, immersive environments, and promoting interaction of users. In contrast, Dolata and Schwabe [11] carried out a fully grey literature review. They reviewed 273 unique newspapers and magazines published in English between 1995 and 2022. For the authors, the construction of the Metaverse occurs in a broader social, technological, organizational, political, and cultural context. They stated that there are multiple metaphors and explanations coexisting simultaneously. Definitions are influenced by the following perspectives: ontological, differential (comparisons with other phenomena), structural (constituents and relationships), and capabilities (what is possible within the Metaverse). The results revealed that social groups are relevant in shaping the meaning and development of the Metaverse; groups include producers (i.e., big tech companies, game producers), users (individuals and retail/entertainment firms), and advocates (investors and governments).

The concept of the Metaverse has been the subject of extensive exploration and definition in the literature. However, most of the studies conducted so far have been focused on academic and technical sources. This has resulted in a need for more research that examines the understanding of the Metaverse using sources that are closer to the industry. Given the fast-growing interest in the Metaverse and its potential applications, it is crucial to have a better understanding of the different perspectives surrounding it. It is worth noting that although Dolata and Schwabe [11] have examined practitioners' perspectives from grey literature, their data sources brought very different views. The authors did not

filter the Metaverse definitions by groups of professionals or ordinary people which resulted in a broad definition. Therefore, our study aims to address this lack of a more focused viewpoint by concentrating effort on getting evidence about the understanding of Metaverse solely from the industry's perspective. Our study addresses this knowledge gap by exploring the insiders' view of the Metaverse.

3 Research Method

Considering the gap in exploring the Metaverse definition from the perspectives of professionals, we decided to conduct an analysis of a specific type of grey literature - podcasts. Our study focused on examining the perspective of practitioners who are actively working in the fields shaping the Metaverse.

Grey literature corresponds to content that is not published in peer-reviewed traditional sources such as academic journals or conferences [12]. It is available in various sources (e.g., technical reports, theses, dissertations, audio and video media, patents). Grey literature content often is produced by professionals who report their practical experience [12]. It has been adopted as a source of valuable information in Software Engineering research as can be seen in [15,17].

Garousi et al. [12] provide a set of questions that support the decision on adopting or not the grey literature as a research source (Table 1). The authors recommend the use of Grey Literature Review (GLR) in the case at least one question has the answer "yes". Taking into account our goal of exploring the Metaverse definition, we have five "yes" answers out of the seven questions.

Table 1. QA to decide whether we should use the GL in our work.

Questions (based on Garousi et al. [12])	Our answers
(1) Is the subject "complex" and not solvable by considering only the formal literature?	Yes
(2) Is there a lack of volume or quality of evidence or a lack of consensus on outcome measurement in the formal literature?	Yes
(3) Is the contextual information important to the subject under study?	No
(4) Is it the goal to validate or corroborate scientific outcomes with practical experiences?	No
(5) Is it the goal to challenge assumptions or falsify results from practice using academic research or vice versa?	Yes
(6) Would a synthesis of insights and evidence from the industrial and academic community be useful to one or even both communities?	Yes
(7) Is there a large volume of practitioner sources indicating high practitioner interest in a topic?	Yes

Considering the relevance of examining the grey literature, we analyzed the perspectives of insider professionals from 10 episodes of a podcast entitled "Into

the Metaverse”². We selected this podcast series because it is from Bloomberg, a well-known broadcaster, and primarily focuses on discussing ‘what is metaverse’ from the perspectives of practitioners who were actively involved with Metaverse. In the following sections, we discuss the data preparation and analysis in detail.

3.1 Data Preparation

The podcast series conducted the interviews from 2021 to 2022 and consisted of 12 episodes and one teaser. Each episode lasted from about 30 min to one hour. We selected 10 out of the 12 episodes for investigation and two episodes were excluded from our sample due to they did not feature external interviewees. In each of the 10 selected episodes, an insider - a professional from different industry sectors (e.g., gaming, business) who is active in the Metaverse arena - was interviewed, providing insights into the conception of the Metaverse. Table 2 shows the title of the selected episodes and the professionals interviewed.

Table 2. Selected Episodes from the podcast series.

Ep.	Podcast title	Interviewee’s role/company
1	Developments, Investments & Experiences in the Metaverse	CEO of SuperSocial
2	Brand Strategies With Cathy Hackl	Leading strategist
3	Building the Metaverse with Marc Petit of Epic Games	General Manager of Epic Games
4	The Creator Economy & the Metaverse With Joost van Dreunen	CEO and co-founder of SuperData
5	The Web 3 Distributed Metaverse with Ryan Gill	CEO of Crucible and Managing Director of the Open Meta Foundation.
6	Roblox’s Growth Opportunity With Chief Business Officer, Craig Donato	Chief Officer of ROBLOX
7	the Metaverse ETF Boom With Mario Stefanidis	Vice president of research from Round Hill investments
8	Blockchain-Enabled Virtual Worlds With Ubisoft’s Nicolas Pouard	Vice President of Ubisoft’s Innovation Lab
9	Into the Omniverse With Nvidia’s Rev Lebedarian	Vice President of Omniverse and Simulation Technology at NVIDIA
10	the Metaverse ETF Boom is No Virtual Reality	ETF Analyst at Bloomberg

As the episodes were in audio format, we transcribed them into a textual format for data analysis. We employed *Whisper*, an open-source³ tool for audio-to-text transcription. Developed by OpenAI, Whisper is an Automatic Speech Recognition (ASR) tool supporting multilingual and multitask⁴, and having an error rate of 3.52% for audios available in the English language [13]. We implemented a Python script coding to use Whisper and get the transcribed texts. A total of 7 h and 45 min of podcast audio resulted in for analysis.

² <https://open.spotify.com/show/7q70azyk47FnPHnCDWuLc7>.

³ <https://github.com/openai/whisper>.

⁴ <https://openai.com/research/whisper>.

3.2 Data Analysis

Taking into account the 135 pages of transcribed text, we conducted a thematic analysis in four steps following the coding technique illustrated in Fig. 1. Open coding technique is a procedure for qualitative data analysis involving decomposing raw data into smaller segments, referred to as codes [14]. The generated codes aim to descriptively and objectively represent the information available in the chunk of text to facilitate subsequent data organization, interpretation, and analysis [14]. We adopted Atlas.ti⁵ tool for the coding process. It is a popular software tool to assist researchers in qualitative data analysis.

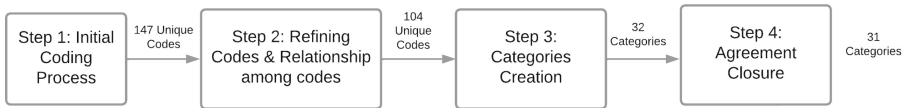


Fig. 1. Data analysis process.

Four researchers participated in the data analysis (see Fig. 1), hereinafter referred to as R1, R2, R3, and R4. R1 and R2 are master students with 2+ years of experience in software engineering. R3 and R4 are senior researchers with 15+ years of experience in qualitative research in software engineering. In the first step, R1 guided their analysis of each podcast episode by searching for evidence that answered the question “What is Metaverse?” as soon as R1 found some chunk of text related to the question, a code was assigned to it. After that, R1 proceeded with a review of the codes to identify codes with substantial similarities, leading to the creation, removal, or merging of certain codes. This step produced 147 initial codes (see Step 1 in Fig. 1). Subsequently, R2 evaluated the codes assigned to the text and the respective code definitions. In Step 2, R1 and R2 held a consensus meeting to consolidate the open coding results, resulting in 104 remaining codes (see Step 2 in Fig. 1).

Before the start of Step 3, R1 reevaluated the podcast episodes and codes to identify intersections within the text. Utilizing the snowball sampling technique across the documents, the researcher uncovered relationships among different codes. Additionally, R1 and R2 worked collaboratively to identify these relationships specifically. In the second part of Step 2, they explored the interconnections of the 104 codes. In Step 3, R1 and R2 collectively defined a set of categories in which the codes were systematically organized. During this phase, the 104 codes were categorized into 32 categories. In Step 4, R3 and R4 reviewed the 32 categories, conducting a double-check of the results. After a consensus meeting involving R1, R2, R3, and R4, two categories were merged, resulting in 31 unique categories. Figure 2 provides an illustrative example of data extraction. The final codes and the respective categories were compiled into a spreadsheet⁶

⁵ <https://atlasti.com>.

⁶ The spreadsheet is available at: https://bit.ly/metaverse_spreadsheet.

After examining the categories, we arranged the 31 categories into three groups that represent *the enabling factors* that will make the Metaverse a reality, *the main characteristics* that the Metaverse presents, and *the impact* that the Metaverse will produce in the world. In the following section, we will focus on *the main characteristics* group, which answers the RQ posted in the Introduction section.

And then the second piece, which is potentially the most transferring and there's so much to unpack on that point is we're going to manifest ourselves into the Metaverse as humans and living in the Metaverse as an avatar. And that avatar doesn't even have to be one avatar. It could be many, many, many avatars. And you and I, Matt, you can find yourself expressing yourself as individual through a wide area of avatars that maybe fit to different type of experiences. Maybe when we want to experience music in the Metaverse, we're going to use, you know, avatar-y and when we want to be more of a sportsman or play sports in the Metaverse, we're going to be avatar-x.

source: Episode 1
category: avatar-identity
code assigned: living in the Metaverse as an Avatar.

Fig. 2. Example of coding result.

4 Results

Table 3 shows the categories of *the main characteristics* of Metaverse, their sub-categories, and the episodes that contained evidence for the categories. In the following sections, each category is presented in detail.

4.1 Metaverse Technology Capabilities

This category encompasses several key technology capabilities that characterize the Metaverse meaning (i.e., *what Metaverse is*) according to the interviewed professionals. It is composed of four sub-categories which are described in the paragraphs below.

Virtual realm: it describes the digital environment where individuals can interact, explore, and engage within the Metaverse, blurring the boundaries between the physical and digital realms. As the vice president of Omniverse and Simulation Technology at NVIDIA declared, “*we need to assemble a virtual world.*” The CEO and co-founder of SuperData is more cautious: “*Virtual reality is something that we’ve seen every decade that comes back and then it becomes nothing and then it comes back again. You know, and it’s always in the future. It’s always this perfect relationship, this perfect technology. And I think the Metaverse is similar...*”. Independently of the terminology, i.e., virtual environment, virtual reality, or virtual world, the interviewed professionals agreed that it is one of the essential aspects of the Metaverse.

Avatar identity: it captures the concept of individuals representing themselves with digital avatars in the Metaverse, allowing for personal expression and adaptation based on context and experiences through multiple avatars. Living in the Metaverse as an avatar or multiple avatars makes it a place to manifest

Table 3. The main characteristics categories

High-Level Category	Sub-categories	Source Ep.
Metaverse technology Capabilities	Virtual Realm	1, 2, 3, 4, 7, 9 and 10.
	Avatar Identity	1, 5, 6 and 7.
	3D representation	3 and 4.
	Integrated simulation and interconnectivity	2, 4, 7 and 9.
Metaverse Infrastructure characteristics	Decentralized	1, 2, 3, 5, 6, 7 and 8.
	Open platform that support persistence and consistency	1, 3, 5 and 8.
	Upcalable	1, 2, 4, 5, 6, 7 and 9.
	Device agnostic	1, 2 and 3.
	Real-time interoperable	1, 3, 4, 6, 7, 8, 9 and 10.
Social/economic aspects	Immersive environment	2, 3, 4, 5, 6, 7 and 9.
	Gaming as primary interaction	1, 2, 3, 4, 5, 7, 8 and 10.
	Co-shaped by both tech and non-tech communities	3, 4, 5, 6, 7, 8 and 9.
	Global economic infrastructure	6 and 8.
	Futuristic temporarility	1, 2, 4, 5, 7 and 10.

oneself. According to the CEO of SuperSocial, avatars are key tools in the Metaverse experience, capable of enabling different types of experiences depending on the avatar type. For him, it is *“potentially the most transferrring and there’s so much to unpack on that point is we’re going to manifest ourselves into the Metaverse as humans and living in the Metaverse as an avatar. And that avatar doesn’t even have to be one avatar. It could be many, many, many avatars.”* The Vice President of Research from Round Hill Investments shares the same line of thoughts, suggesting that the possibility of avatars is a factor for the decision of interacting in these spaces: *“The reason that consumers want to interact in these spaces is this concept of expressing yourself with your avatar. Digital self-expression is, I like to call it like that. The avatar economy is what the younger generation likes to call it.”*

3D representation: this sub-category refers to the need to include three-dimensional digital objects and environments within the Metaverse. The professionals interviewed in the podcasts believe that 3D is essential for representing the Metaverse, *“whether we like it or not”* (mentioned the CEO and co-founder of SuperData). The general manager of Epic Games thinks that the Metaverse *“is going to be born out of the revolution around the World Time 3D. As World Time 3D becomes a mainstream medium, it becomes easy to capture 3D and everybody can consume interactive 3D content, because they have a powerful device or it’s streamed from the cloud.”*

Integrated simulation and inter-connectivity: quite a few professionals believe that the Metaverse needs a holistic approach to combining software and hardware elements. As the vice president of Omniverse and Simulation Technology at NVIDIA explains, *“our unique contribution to this thing we’re calling the Metaverse and the future of computing is powering all of the simulation necessary to do this. That’s not just a hardware problem. It’s a combination of software and hardware problems.”* The accurate modeling of physics-based simulations is needed to ensure the faithful representation of the laws of physics and the interactions of objects within the virtual environment. It also highlighted the

concept of bridging the gap between the physical and digital worlds, involving a seamless connection and interaction between tangible reality and virtual spaces.

4.2 Metaverse Infrastructure Characteristics

Claimed as the new Internet, there are key characteristics that the active players in the arena believe that the Metaverse infrastructure should embrace in order to exist and function on a global scale.

Decentralized: by being decentralized, the Metaverse provides enhanced security, transparency, and decentralized control over data management. For the Director of the Open Meta Foundation, decentralized technology is non-negotiable: *“the Metaverse is just a phase of the Internet that we’re kind of going through right now. To me, there are some non-negotiable [things]. I believe that it needs to be decentralized. I think the only way to have Web3 is through decentralization”*. Blockchain is at the very center of decentralized technology. Even though some believe that it is an optional solution, it is considered necessary to foster an environment where users and developers have the freedom to integrate blockchain technology into their Metaverse experience. According to the Vice President of Ubisoft, blockchain represents the core feature of Metaverse: *“I’m very, very bullish about that. I’m pretty sure that without blockchain there is no Metaverse... The idea is, with decentralization, you share the infrastructure, then you are creating trust [in the environment] and from that trust, you can create this representation of the new value [and] we all share, and you can distribute this more fairly”*.

Open platform that ensures persistence and consistency: There will be challenges in maintaining soundness among different virtual worlds in the Metaverse, ensuring that they align with shared standards and guidelines. Standardization is needed to provide consistency, persistence, and compatibility across platforms and applications. An open platform ensures coherence, continuity, and longevity within the virtual worlds of the Metaverse, as the Leading strategist interviewed in Episode 2 (see Table 2) claimed: *“we’re going to have to invent a new infrastructure, [and we need to] manage that openness”*. This may not be easy, as the CEO of Crucible and Managing Director of the Open Meta Foundation commented: *“the Metaverse is emerging as the next big technology platform as I like to say on this podcast. That’s why Apple and Epic are fighting now. Epic talks about open standards and being an open Metaverse platform”*.

Upscalable: the Metaverse will be a large-scale environment that can be scalable. The professionals point out that *“this thing we’re calling the Metaverse, or Web3, or whatever it ends up being called... the scale of it and the exact shape and feeling it, we can’t predict. But one thing I think we can be sure is that it’s going to be bigger than anything we’ve ever known”*, mentioned the vice president of Omniverse and Simulation Technology at NVIDIA. Therefore, the Metaverse needs to be *“upscaleable”*, possible for millions of concurrent players, and support the distribution of human behavior over the internet and large-scale simulations. As the CEO of SuperSocial envisioned, *“the dream of the Metaverse is of course that not couple hundred people can experience a concert of robots [...] actually*

it's millions of people congregating in one place at one single point in time to experience something together".

Device agnostic: The advocates of the Metaverse believe that it is not tied to any specific device, and it can be accessed and experienced across various platforms and technologies. Both the CEO of SuperSocial and the General Manager of Epic Games commented that "*obviously Metaverse experiences are going to be accessible through any device. And so the question of what platform people are going to consume information or experiences on... it doesn't really matter, because we're going to be able to access those experiences from any device*". This statement came up "*to sort of demystify*" that we're going to access the Metaverse from one single device such as VR glasses.

Real-time interoperable: the Metaverse is "*going to be an interoperable synchronous persistent series of virtual space*", stated the vice president of research from Round Hill Investment. It is real-time and always on, featuring user synchronization and responsive feedback. The vice president of Omniverse and Simulation Technology at NVIDIA claimed that "*for the Metaverse to exist, there must be interoperability*". The insiders believed that these characteristics provide dynamic, interactive, and immersive experiences to users.

4.3 Social/economic Aspects

This category represents the essential non-technical characteristics of the Metaverse.

Immersive environment: it describes the quality of experiences within the Metaverse that support the deep engagement of users' senses, creating a sense of presence and realism through advanced technologies, high fidelity, and spatial interactions. For some, the capacity of being immersiveness is one decisive characteristic of the Metaverse and is a "*kind of gate to its adoption*" of it. However, when the virtual and physical worlds become indistinguishable, the Metaverse can be a way for some users to escape reality, which may bring negative consequences to their personal and social life and well-being.

Gaming as primary interaction: gaming is a central focus into the Metaverse, playing a pivotal role in shaping and popularizing virtual worlds. For the vice president of Ubisoft's Innovation Lab, "*at least in the foreseeable future, the Metaverse is still going to be predominantly about gaming*". The gaming companies have built a massive user base and they are investing in gaming to have the content to support their Metaverse efforts. Therefore, gaming will continue to be a key driver in the early stages of the Metaverse maturity, serving as a way to popularize Metaverse immersion.

Co-shaped by both tech and non-tech communities: this sub-category emphasizes the understanding that the Metaverse is co-created and shaped by both its developers & community and users. It is a community space that motivates various types of collaboration and co-creation of innovative products, services, or experiences. As the vice president of research from Round Hill Investment claimed, "*to me, that's what's sort of really, really exciting about the*

Metaverse as a place for human experience, human interaction, playing, working, doing things together”.

Global economic infrastructure: it encompasses the elements that contribute to the establishment and operation of an economic system. The Metaverse insiders recognized the importance of a robust economy that allows users to engage in buying, selling, and earning which can add value to the virtual experience. As the Chief Officer of ROBLOX explained, “*all these experiences in this universe are integrated with a common fabric. And that fabric has a couple of different dimensions to it. You know, it has a common identity framework, you’re the same person. It has a social graph, right? I go around with my friends. It has an economic [ecosystem]. I’m able to buy, sell, and make a living across these different experiences*”. The demand for negotiating things requests a common digital currency or monetary system to facilitate transactions and economic activities on a global scale, as the vice president of Ubisoft’s Innovation Lab argued, “*without global currency, you don’t have a Metaverse... Gold [used as currency] was the standard for all monetary systems, pre-World War One, Bitcoin could become that new standard*”.

Futuristic temporality: For some of the interviewees, the Metaverse concept has a temporal dimension that encompasses the understanding that the Metaverse “*is not something that’s going to be realized overnight. It’s going to be probably a decade or more until there is actually a Metaverse in place*”, mentioned the CEO and co-founder of SuperData. There is also the opinion that the Metaverse is not only a virtual world or a set of technologies. It is “*a point in time*” when people stop making the distinction between the virtual worlds and the physical ones.

5 Discussion

The analysis of the 10 podcast episodes supported us to answer our RQ (*What is the definition of the Metaverse from the perspectives of the involved practitioners?*). First, our findings confirmed that the sole definition of Metaverse can hardly be achieved due to the complexity and multifaceted of the themes that compose it. This perception is aligned with the discussions previously presented in our related work (see Sect. 2). Unlike the related work, we could see from our results that there are high-level groups that provide a viewpoint on *the enabling factors* to become the Metaverse a reality, and *the main characteristics* of the Metaverse and *the impact* that the Metaverse will produce in the world. In this paper, we concentrated on discussing the *the main characteristics* group which covers three categories.

Taking into account the three categories presented in this paper, we can see that 3D representation, avatar identity, immersive environment, and virtual realm, i.e., elements of *metaverse technology capabilities category*, have already appeared across the related work [1–3, 7]. This similar result restated these elements as core features of the Metaverse that show a consensus from the definitions presented in other works. Although the use of avatars has been found in

the literature recurrently, our results unfold a new expression for defining such practice: *digital self-expression*. It represents a new form of showing how people see themselves from a picture they created. Nonetheless, the results revealed that the professionals are concerned about connecting users to the Metaverse considering the endeavor of integrating a complex environment with different technologies and the available connectivity (i.e., integrated simulation and inter-connectivity, a new category uncovered in our work). Our result emphasizes the importance of having properly interconnected devices and software to provide a seamless simulation. Park and Kim [10] provided a similar discussion but as a simpler view of the relationship between software and hardware.

Considering the *metaverse infra-structure characteristics* category, the results revealed that most of the elements have been discussed in the literature [2, 7, 9]. However, we could see that the discussion about the scalability of the Metaverse environment (i.e., upscalable sub-category) attained new concerns about the sharing of the Metaverse infrastructure and the value that this practice could bring to the trust of using the environment. The *device agnostic* was also another new element uncovered in our study that gives the perspective that there are various means of accessing the Metaverse that involve multiple device types.

Finally, the results showed an evolution in the discussion about the *social/economic aspects* related to the Metaverse. Elements such as the immersive environment, interaction from games, global economic infrastructure and the participation of tech and non-tech communities in the co-shaping of the Metaverse have been addressed in the literature [1–3, 9, 11]. Our results reaffirmed the tendency for discussions about these elements to mature within the industrial context. However, the futurist temporality sub-category emerged from the results as a futuristic concept that professionals will strive to understand. It may represent a rupture of the viewpoint of online communication due to it can make it difficult the distinguish between the interaction that happens in the physical world and the ones that occurs in the virtual environment. This perspective triggers an ethical and crucial discussion on the direction that society will evolve and the relationship among people.

Although our study brings contributions to the exploration of the Metaverse definition, we understand that it has some limitations. First, we have the conscious that the Metaverse is an evolving concept and defining it solely based on insights from professionals may not encompass all characteristics and future developments. Even though the interviewed professionals in the 10 podcast episodes come from different types of companies and assume various roles, the sample size is relatively small. Therefore, the findings can not be generalized as the shared understanding by all professionals working in the Metaverse-related fields. More interviews of professionals, either by collecting more grey literature or by conducting interviews directly with them, will increase the generalizability of the findings obtained in this study.

6 Conclusion

In this work, we presented a study that explored the definition of the Metaverse from the perspective of insiders actively working in this field. To achieve this, we analyzed 10 podcasts, i.e., grey literature, which contains interviews with professionals from different companies. Our main result was the identification of the essential characteristics of the Metaverse concept that we classified into three categories, i.e., the Metaverse technology capabilities, infrastructure characteristics, and social and economic aspects. Each category presented elements which supported us to discuss different elements that impact the Metaverse definition.

As a contribution, we restated some important elements requested to the Metaverse definition that have been covered from the literature as well as unfolded new ones. We could see that some common elements that appeared in the literature, e.g., the use of avatars, are now recognized as a way for users to express their view of themselves. Besides, the adoption of multiple devices, infrastructure sharing, and the recognition of real and virtual worlds are concerns of industrial professionals that deserve more discussions. In future work, we intend to explore further the other main groups of categories that we have found in our study. These categories certainly can expand the understanding of the Metaverse phenomenon.

Acknowledgments. We thank Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil (CNPq grant 309497/2022-1) for the partial financial support. We also thank Alessandro Aneghi for supporting us in the data analysis process.

References

1. Weinberger, M.: What Is metaverse?-A definition based on qualitative meta-synthesis. *Future Internet* **14**, 310 (2022). <https://doi.org/10.3390/fi14110310>
2. Almoqbel, M.Y., Naderi, A., Wohn, D.Y., Goyal, N.: the metaverse: a systematic literature review to map scholarly definitions. In: *Companion Computer Supported Cooperative Work and Social Computing*, pp. 80–84. ACM, Virtual Event Taiwan (2022). <https://doi.org/10.1145/3500868.3559448>
3. Ng, D.T.K.: What is the metaverse? Definitions, technologies and the community of inquiry. *AJET*. **38**, 190–205 (2022). <https://doi.org/10.14742/ajet.7945>
4. Zhang, X., Chen, Y., Hu, L., Wang, Y.: the metaverse in education: definition, framework, features, potential applications, challenges, and future research topics. *Front. Psychol.* **13**, 1016300 (2022). <https://doi.org/10.3389/fpsyg.2022.1016300>
5. Abu-Salih, B.: MetaOntology: toward developing an ontology for the metaverse. *Front. Big Data.* **5**, 998648 (2022). <https://doi.org/10.3389/fdata.2022.998648>
6. Hwang, G.-J., Chien, S.-Y.: Definition, roles, and potential research issues of the metaverse in education: an artificial intelligence perspective. *Comput. Educ. Artif. Intell.* **3**, 100082 (2022). <https://doi.org/10.1016/j.caeai.2022.100082>
7. Ritterbusch, G.D., Teichmann, M.R.: Defining the metaverse: a systematic literature review. *IEEE Access.* **11**, 12368–12377 (2023). <https://doi.org/10.1109/ACCESS.2023.3241809>
8. Gökçe Narin, N.A.: Content analysis of the metaverse articles. *J. Metaverse*, **1**(1), 17–24 (2021). <https://dergipark.org.tr/en/pub/jmv/issue/67581/1051382>

9. Chen, T., Zhou, H., Yang, H., Liu, S.: A review of research on metaverse defining taxonomy and adaptive architecture. In: 2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), pp. 960–965. IEEE, Chengdu, China (2022). <https://doi.org/10.1109/PRAI55851.2022.9904076>
10. Park, S.-M., Kim, Y.-G.: A metaverse: taxonomy, components, applications, and open challenges. *IEEE Access* **10**, 4209–4251 (2022). <https://doi.org/10.1109/ACCESS.2021.3140175>
11. Dolata, M., Schwabe, G.: What is the metaverse and who seeks to define it? Mapping the site of social construction. *J. Inf. Technol.* **38**, 239–266 (2023). <https://doi.org/10.1177/02683962231159927>
12. Garousi, V., Felderer, M., Mäntylä, M.V.: Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Inf. Softw. Technol.* **106**, 101–121 (2019). <https://doi.org/10.1016/j.infsof.2018.09.006>
13. Radford, A., Kim, J.W., Xu, T., Brockman, G., McLeavey, C., Sutskever, I.: Robust speech recognition via large-scale weak supervision (2022). <https://doi.org/10.48550/ARXIV.2212.04356>
14. Saldaña, J.: *The Coding Manual for Qualitative Researchers*, 2nd edn. SAGE, Los Angeles (2013)
15. Melegati, J., Guerra, E., Wang, X.: Understanding hypotheses engineering in software startups through a gray literature review. *Inf. Softw. Technol.* **133**, 106465 (2021). <https://doi.org/10.1016/j.infsof.2020.106465>
16. Wang, H., et al.: A survey on the metaverse: the state-of-the-art, technologies, applications, and challenges. *IEEE Internet Things J.* **10**, 14671–14688 (2023). <https://doi.org/10.1109/JIOT.2023.3278329>
17. Zaina, L.A.M., Fortes, R.P.M., Casadei, V., Nozaki, L.S., Paiva, D.M.B.: Preventing accessibility barriers: guidelines for using user interface design patterns in mobile applications. *J. Syst. Softw.* **186**, 111213 (2022). <https://doi.org/10.1016/j.jss.2021.111213>




Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Carbon Footprint Calculations for a Software Company – Adapting GHG Protocol Scopes 1, 2 and 3 to the Software Industry

Antti Sipilä¹ , Laura Partanen² , and Jari Porras^{2,3,4} 

¹ TIEKE, Helsinki, Finland

² LUT University, Lappeenranta, Finland

laura.partanene@lut.fi

³ Aalto University, Helsinki, Finland

⁴ University of Huddersfield, Huddersfield, UK

Abstract. Through non-financial reporting, such as CSRD, carbon footprint calculations are becoming mandatory in the software industry. The golden standard for reporting CO₂ emissions is based on the Greenhouse Gas (GHG) Protocol and its scopes 1, 2, and 3. However, as a producer of purely digital products, the software industry differs from traditional industries in its carbon footprint. The software industry value chain relies heavily on an infrastructure that can contribute most of its emissions. It has been recognized that there is a need for an industry-customized carbon emissions model that considers the software industry's peculiarities. The primary goal of this study is to define the main sources of climate impacts in the software industry and propose a model of the GHG Protocol adaptation to software companies. This research has been done in our Green ICT project and is based on interviews done in that project. The data for this research was collected from five software companies with different demographics and business models. The interviews, with a total amount of 14, were conducted between November 2022 and March 2023 during a service design process of an automated tool that facilitates green transition in software companies. The analysis of the interviews was supplemented with the results from four multi-stakeholder workshops conducted during the service design process, as well as with the analysis of a series of webinars around the topic. As a result of the study, the Software Company Scopes model for the primary sources of greenhouse gas emissions in the software company and its value chain was created, and the GHG Protocol was tailored to the needs of the software industry. Thus, considering its industry-specific peculiarities, we may conclude that the GHG Protocol can be applied to the software industry.

Keywords: Software Company · Greenhouse Gas · Reporting

1 Introduction

Within the last 15 years, since the publication of Global eSustainability Initiative's (GeSI) SMART2020 report in 2008 [1], awareness about the ICT industry's carbon handprint and footprint has increased. According to the latest GeSI SMARTer2030 report

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBP 500, pp. 442–455, 2024.

https://doi.org/10.1007/978-3-031-53227-6_31

[2] ICT has a large handprint potential of about 12,08 Gt CO₂e, while the footprint is about one-tenth of this, 1,25 Gt CO₂e. While the handprint potential is substantial, we can not ignore the footprint, as, according to the report, it is the fastest growing of all industries, projected to triple between 2015 and 2025.

As Freitag et al. [3] state, the ICT sector has become a significant factor in global carbon emissions. It is estimated in their study that the ICT sector creates 2.1–3.9% of global greenhouse gas emissions. It is self-evident that this is a subject that needs to be noticed if we want to achieve the objectives of the Paris Agreement¹ to “hold ‘the increase in the global average temperature to well below 2 °C above pre-industrial levels’ and pursue efforts ‘to limit the temperature increase to 1.5 °C above pre-industrial levels’.” The EU executes this with the initiative of the European Green Deal², which shows the path for Europe to be climate-neutral by the year 2050. EU is controlling this objective through the European Climate Law³. Currently, EU directive NFRD EU/2014/95⁴ determines the need for large public interest entities with over 500 employees, such as banks, insurance companies, and bigger listed companies, to make “a non-financial statement containing information to the extent necessary for an understanding of the undertaking’s development, performance, position and impact of its activity, relating to, as a minimum, environmental, social and employee matters, respect for human rights, anti-corruption and bribery matters.” EU Directive 2022/2464⁵ of corporate sustainability reporting “modernizes and strengthens the rules concerning the social and environmental information that companies must report. A broader set of large companies, as well as listed SMEs, will now be required to report on sustainability.” The new directive will be implemented in reporting for the first time for the financial year 2024. The reporting should be done according to European Sustainability Reporting Standards (ESRS)⁶. The company-specific Greenhouse gas emissions are to be reported within the scopes one, two, and three adopted from the GHG Protocol [4]. In short, scope one emissions are direct emissions from the company operations, scope two emissions are formed from the energy used in the company, and scope three emissions include all the indirect emissions in the value chain, in both up and downstream activities. This may become a challenge for software companies since their business operations produce immaterial products. This will be further discussed in Sect. 2.2.

1.1 Green ICT Ecosystem Project

This research is based on work done in the Finnish Green ICT ecosystem -project⁷. The project aimed to increase the environmental awareness of Finnish ICT companies and build an ecosystem around the topic of Green ICT in the Uusimaa region. The

¹ <https://unfccc.int/process-and-meetings/the-paris-agreement>.

² https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/climate-action-and-green-deal_en.

³ https://climate.ec.europa.eu/eu-action/european-climate-law_en.

⁴ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0095>.

⁵ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022L2464>.

⁶ <https://www.efrag.org/lab3>.

⁷ <https://tieke.fi/en/projects/green-ict-project/>

project provided webinars, online workshops, and published guides to both procurers and producers of ICT products and services. The concrete outcome besides the guides was a web-based self-assessment tool for organizations to evaluate their level of climate and environment-neutral actions and to provide a base for their development plan. In the development of the tool, a service design process was utilized. The service design process used the double diamond model [5], a widely used method in service design processes, which will be presented with a wider lens in Sect. 3.1. The design process has been used in this research as a basis for the development of the software industry-specific carbon emission model, which will then help the software companies report their carbon emissions.

1.2 Objective of the Study

The objective of this study is to define the GHG components in scopes one, two, and three for software companies. By providing these software-specific components, reporting their carbon emissions becomes a bit easier. The practical need from software companies and our project objective led to our research question: What should software companies report within scopes 1, 2, and 3?

By providing an answer to this research question through design science research, we aim to contribute to the EU-level objective of carbon emission reporting in every industry sector.

2 Background

In this Background section, we present the Greenhouse Gas (GHG) Protocol that forms a basis for our model. We also describe the software industry emissions on a general level and the challenges the software industry may have while using the general GHG Protocol.

2.1 Greenhouse Gas Protocol

With the increase in awareness of the negative effects of human activity on the climate, mainly particle pollution, international bodies and forums have started preparing mitigation measures. This raised the issue of defining and calculating the emissions to understand the challenge clearly. As with all emerging fields, varying methods of emission calculations arose early on, and standardization became a necessity as the results were about as comparable as apples and bananas. This standard needed to address factors such as emission equivalency, comparability, assigning of responsibility, and sustainability reporting usability.

Greenhouse Gas Protocol [4] has emerged as the most popular and is widely regarded as the golden standard method of emission calculations. International Standardisation Organisation's (ISO) standard for carbon emission calculations, ISO 14064 [6] is compatible with the GHG Protocol, and it is being used by, for example, Global Reporting Initiative (GRI)⁸ and Science Based Targets Indicators (SBTi)⁹.

⁸ <https://www.globalreporting.org/>

⁹ <https://sciencebasedtargets.org/>

GHG Protocol calculates emissions as carbon dioxide equivalent (CO_2e), in which all different greenhouse gas emissions can be measured. The equivalency is calculated in relation to each emission's atmospheric warming potential in comparison to carbon dioxide for conversion into a comparable metric. The protocol divides emissions into three scopes according to the source (see Fig. 1) [4]. Figure 1 presents these scopes, as depicted by the Environmental Protection Agency of the United States. It represents these emissions in the three scopes of the GHG Protocol, divided between the upstream and downstream activities of the reporting company. Upstream of the value chain pertains to anything that is procured by the reporting company, and downstream pertains to anything produced and sold by the reporting company. Scope one contains the direct emissions from the operations of the measured company, such as equipment and office. Scope two pertains to indirect emissions that are caused by energy usage of the reporting company and are caused in upstream of the value chain. Scope three emissions are divided into the upstream and downstream activities. Upstream Scope three emissions are caused by different products and services used by the reporting company, including, for example, diverse emissions from employee commutes and purchased services. Downstream Scope three emissions pertain to activities conducted in advertising, sales, distribution, and usage of the reporting company's products, in addition to investments made and financial assets held by the reporting company.

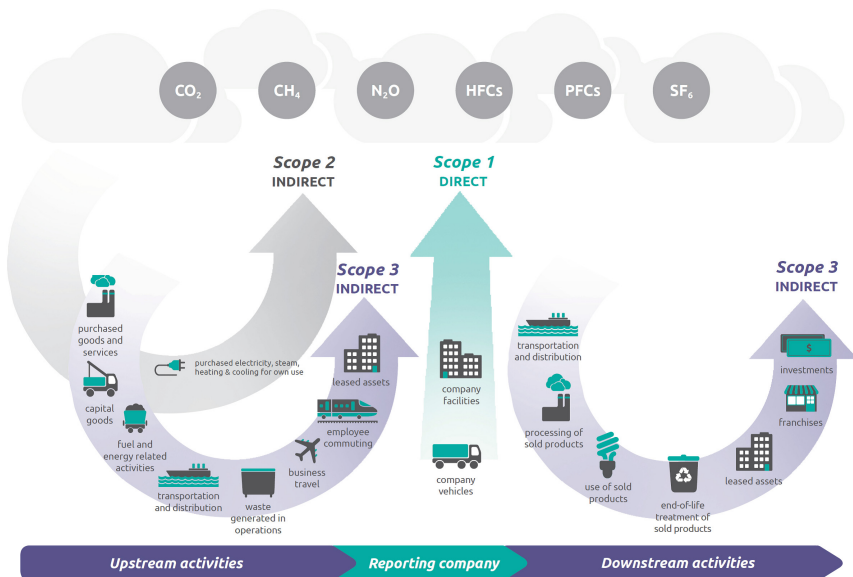


Fig. 1. Greenhouse Gas Protocol Scope Definitions (EPA) [7]

According to A Corporate Accounting and Reporting Standard [8], these scopes are defined as follows:

- **Scope one** emissions represent the direct emissions “owned or controlled by the company” from the business operations of the company in question, such as its equipment and offices.
- **Scope two** emissions represent the indirect emissions from the generation of the electricity used in the company.
- **Scope three** expands the emissions to the company’s value chain and considers both the upstream and downstream activities, such as sub-contractors and clients.

2.2 Challenges of the Software Industry for Emissions Reporting

Adopting the GHG Protocol for specific industries requires identifying the relevant business operations and their effects. Software companies are a special case in this regard, as their products are digital instead of physical. On the other hand, these products are dependent on physical hardware infrastructure, which means they require electricity and thus produce emissions [9]. In addition, modern software uses a client-server model, which runs on a server in a data center environment or a cloud service. This affects the emissions and makes the emissions calculation fuzzy. This is something that has emerged among companies – how should one measure the carbon footprint in such an ecosystem, in which its code and software run on an external data center or services of a third party and are used by another third party? All these factors need to be considered, and decide what of those needs to be calculated.

Software lifecycle can be broadly seen in three stages: the requirement and design phase, the development phase, and the use phase [10–12]. Software is coded to fulfill a specific purpose, whether professional or recreational. In both cases, the purpose defines the requirements that are used in its design [13]. Many of the most relevant decisions that define the software’s climate and environmental impact are made in this first phase of requirements and design [14, 15]. In the development phase, the software is programmed and tested on how well it fulfills these requirements [13].

Digital products are not limited by physical resources and manufacturing, which makes them easily replicable and scalable. Combined with digital distribution, physical media can be bypassed entirely. On the one hand, the non-dependence on physical resources lessens the environmental impact of the products; on the other, the replicability increases the climate impact specifically. This is an issue in downstream scope three emissions.

Another challenge with the software is the variety of client devices used by the end users. These devices have different hardware architectures and energy usage patterns, which raises further challenges in calculating the use phase emissions. This is an issue in downstream scope three emissions.

3 Research Process

In this Research process section, we present the methods used in this study. We also visualize the process of developing the Software Company Scopes model.

3.1 Methods

In this research, we have conducted design science research methodology (DSRM) by Peffers et al. [16]. According to Peffers et al. [16] the design science process consists of six steps. These steps are

1. problem identification and motivation
2. definition of the objectives for a solution
3. design and development
4. demonstration
5. evaluation and
6. communication.

The core of this method is an artifact created during the research process to solve the problem identified in the beginning (Fig. 2). In this study, we present the Software Company Scopes model as an artifact to solve the challenges in the software industry to calculate carbon emissions as presented in Sect. 2.2.

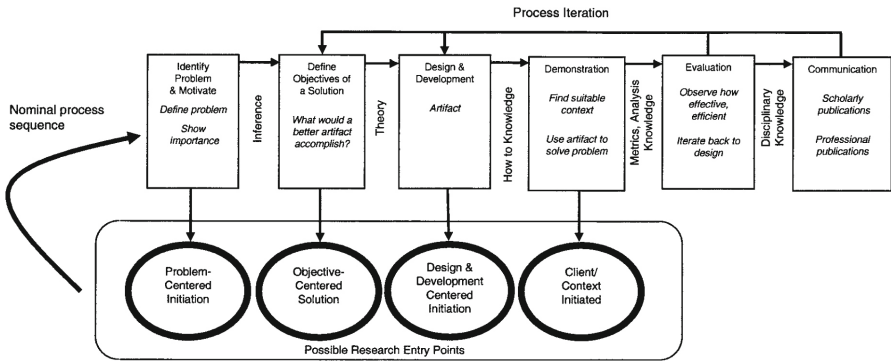


Fig. 2. DSRM Process Model according to Peffers et al. [16]

In this study, we identified the problem (step 1) within the Green ICT project and formed the research question presented in Sect. 1.2. For steps 2–5, we have utilized the double diamond service design process model [5] (Fig. 2) for developing the self-assessment tool described in Sect. 1.1. The double diamond model includes similar components and phases to the DSRM model presented above. The first phase in the double diamond model is understanding, followed by the phase of brainstorming. After these phases, an outcome will be tested and implemented. In this study, the outcome was the web-based self-assessment tool (Fig. 3).

The primary method for data collection used in this study is an interview. Interviews were utilized as expert interviews during the service design process, where there were three rounds of interviews conducted with five different companies from the IT sector. Interviews were executed online via Teams meetings. Participated companies are presented in Table 1. Company E participated only in the first and the second rounds of interviews hence the total number of interviews was 14. The objectives for every round of

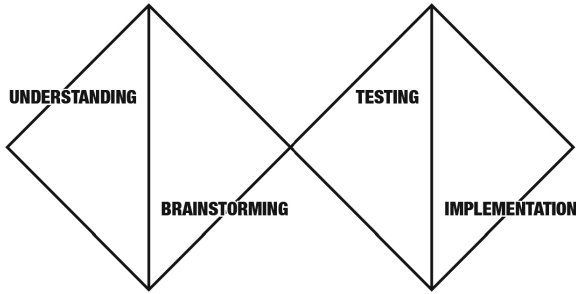


Fig. 3. Double diamond service design model [5]

interviews were different. Objectives and types of interviews follow the double diamond service design process model used in the study and were as follows.

1. Understanding the state of the art in companies and the need for the tool through semi-structured group interviews (companies A-E)
2. Testing the preliminary questions for the tool (companies A-E)
3. A pilot study of the beta version of the tool (companies A-D)

Table 1. Participants in the service design process.

ID	Principal industry (TOL2008 ¹⁰)	Number of employees	Number of participants in interviews	Role of the participants in the company
A	Advertising agencies	10–49	2	Owner/founder, CTO
B	Software design and manufacturing	10–49	3	CEO, Senior Software And UX Designer
C	Software design and manufacturing	10–49	3	CEO, Design lead, Full stack developer
D	Software design and manufacturing	50–249	2	Senior manager, Chief Architect
E	Other professional, scientific and technical activities	under 10	3	CEO, CTO, CIO

3.2 Analysis Process

The analysis of the interviews, which are presented in Sect. 4, was supplemented with an analysis of six webinars and eight ecosystem meetings¹¹ held during the Green ICT

¹⁰ <https://www2.stat.fi/en/luokitukset/toimiala/>

¹¹ <https://tieke.fi/hankkeet/greenictanke/green-ict-tapahtumat/>

project within the time period of October 2021 until August 2023. In the webinars, three companies or organizations represented their work as a business case, product case, or general work in green ICT. These cases included carbon calculation of both software products and SME companies. At the end of the webinar, there was a panel discussion between the participants on the themes of their presentation.

Ecosystem meetings were more varied, and there were discussions and workshops about innovation & research, emission calculations, green coding, green procurement, ICT equipment and its lifetime impact, and sustainable software business models and tools. Analysis of the transcripts from the webinars and ecosystem meetings formed the base information for the questions used in the interview process.

In addition to these webinars, four workshops on the service design process replenished the analysis of interviews. These multi-stakeholder workshops were executed during October and November 2022. The relation between these data collection sets and the structure of the development of the Software Company Scopes model is presented in Fig. 4. With the visualization (Fig. 5), we also present the relation of the GHG Protocol to our model.

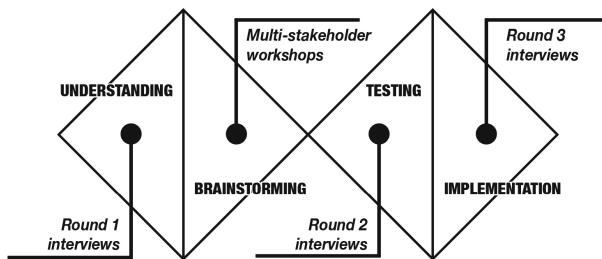


Fig. 4. Steps in this research in relation to the double diamond service design model.

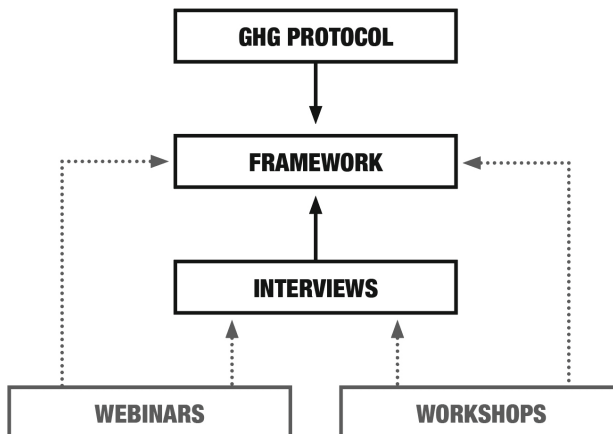


Fig. 5. Visualization of the data collection for the framework.

4 Results

This section presents the software company-specific scopes as a result of our study and the results that led to the model.

4.1 Interviews

The main objective of the first round of interviews was to gain an understanding of the current situation in the companies and the possible challenges they are facing with taking climate and environmental impacts into account in their operations. The main findings from the first round were as follows:

- Information about software-specific climate and environmental impacts is here and there
- What should we as a software company calculate in scopes 1, 2, and 3?
- It is hard to find concrete information or guidance on what to do and how towards more climate-neutral actions
- Regardless of the remote work, there are several offices
- “We produce intellectual property, so it’s hard to have the same way influence on climate issues”
- “A complex thing [climate and environmental impacts of software company] and many things affect another”

From these findings, we generated the analysis:

- There is an obvious need for some concrete guidance on what to include in software companies into the GHG Protocol scopes
- The core of the company requires basically only some facility (office), a computer, and a network.
- The effects of the operations of software companies extend far into subcontracting chains

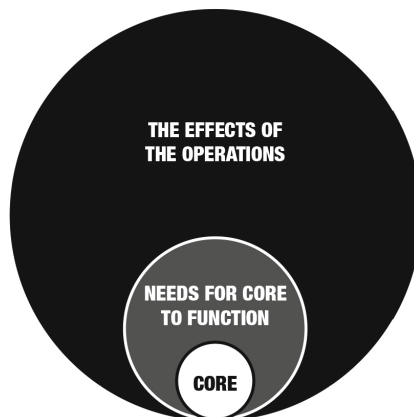


Fig. 6. Visualization of software company functionality

- The layered structure of software companies (see Fig. 6)

Scopes one, two, and three can be directly derived from the image: core functions belong to Scope One, needs for the software company core to function belong to Scope Two, and the effects and operations in subcontracting and distribution chain belong to Scope Three.

4.2 Scopes of a Software Company Framework

After completing the understanding phase of the service design process of the self-assessment tool, we divided the software production process into the following parts based on the analysis of the interviews in the brainstorming and testing phases (see Fig. 2).

1. Organization strategy
2. Software production
 - a. Design
 - b. Coding and testing
 - c. Usage and maintenance
3. Support functions

This division, while somewhat artificial, sheds light on the different Scopes in both upstream and downstream factors and is a useful categorization. In this approach, the decision of how to react to the legislative and public moral pressure is covered in the organization's strategic work. This contains the values, vision, mission, strategy, and action plan of the company. It also includes how the company's staff is informed on how to take climate and environment into account in their work. As the demands pertaining to the company's supply chain are strategic choices, the emission demands from subcontractors are included here.

The practice of how well the Scopes are covered is in the second part, the software production. The first step, design, is the phase where most of the critical decisions concerning the emissions are made [13]. These include architecture choices [17, 18], programming language [19–21], integrated development environment, graphical choices [22], etc. These choices influence both the coding and testing phase and the usage and maintenance phase. As such, it seems to influence many of the scope three emissions in both the downstream and upstream.

The coding and testing phase is the source of Scope One and Two emissions, as it is the main business activity of the company. It is where they use their equipment and offices, and it causes a lot of its direct use of energy. It also includes some Scope Three emissions from the upstream, such as employee commuting.

The usage and maintenance phase is composed mostly of Scope Three emissions from downstream, such as distribution and tech support.

Support functions include the climate and environmental choices made by the company in its everyday operations not directly related to its main business activity. This also includes human resources and marketing. The most important of these are the sourcing of energy, local energy generation, employee training in sustainability competence, and environmental systems present in the offices.

From the division together with the GHG Protocol, we have derived and named the factors to be included in Scopes One, Two, and Three for software companies (Table 2).

Table 2. Factors for a software company to include in Scopes 1, 2, and 3.

SOFTWARE COMPANY					
DIRECT	SCOPE I	OFFICE	Direct emissions related to an office. Includes office building's energy efficiency defines how much energy its heating and AC use. Its location is also a factor in the commuting for employees.		
		VEHICLES	Direct emissions from company's vehicles. Includes company owned vehicles used in its business activities, such as sales and tech support are included here.		
		EQUIPMENT	Direct emissions from equipment. Includes all equipment owned by the company and used on premises are included here. This covers the infrastructure and the user devices.		
INDIRECT	ENERGY	SCOPE II	ELECTRICITY	Emissions from purchased electricity. Includes the source of the electricity used in a company's own business operations.	
			NETWORK	Emissions from network traffic. Includes the amount of traffic and types of connections used in a company's business operations.	
			COOLING	Emissions from AC and equipment cooling. Includes the method of cooling offices and equipment in warmer climates and warm seasons.	
			HEATING	Emissions from space heating. Includes the method of heating offices in colder climates and cold seasons.	
	OTHER	SCOPE III	UPSTREAM	TRAVEL	Emissions from work travel. Includes business travel connected to operations, the method of travel used and amount of travelling done.
				COMMUTING	Emissions from employee commuting to an office. Includes the method of commuting used and the amount of days spent working in the office.
				LOGISTICS	Emissions from material logistics. Includes any transport of equipment and materials by a company's order.
				LEASED ASSETS & AAS	Emissions from leased equipment, vehicles and other assets, as well as those purchased as a service.
				SUPPLIES	Emissions from supply chains of office supplies that are needed in everyday business operations. Includes small supplies and furniture.
				BUILDING AND MANUF. SCOPE I ASSETS	Emissions from the first part of the life cycle of buildings, vehicles and equipment and the relevant supply chains therein.
				WASTE	Emissions from waste processing. Includes all waste produced by any of the company's staff in office.
				SERVERS & CLOUD	Emissions from leased or aaS virtual assets. Includes any third party data centre or cloud service used in the operations of a company.
	OTHER	SCOPE III	DOWNSTREAM	MARKETING	Emissions from marketing activities. Includes any marketing done by the company, whether digital or physical and also third party marketing partners.
				DISTRIBUTION	Emissions from distribution of products. Includes all methods of distribution, whether digital or physical media.
				RETAIL & FRANCHISE	Emissions from retailers and franchise holders. Includes any third party retailer or franchise selling the software.
				TECH SUPPORT	Emissions from tech support activities. Includes all forms of support from automated bots to chat and phone service to support in customer location.
PRODUCTS				Emissions from product use by the customers.	
INVESTMENTS				Emissions from financial assets held by a company.	

As a final result of this study, we have created a Software Company Scopes model similar to GHG Protocol to present the result in an understandable but also comparable form (see Fig. 7).

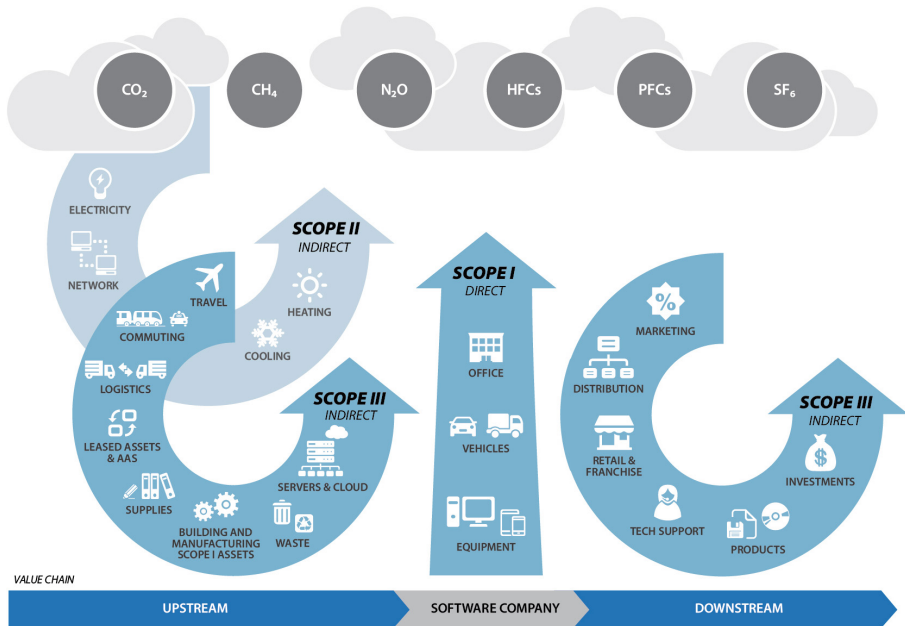


Fig. 7. The Software Company Scopes model presents an overview of scopes and emissions across the value chain of a software company with visualization adopted from the GHG Protocol Corporate Value Chain Accounting Reporting standard [8].

5 Discussion and Conclusion

Software companies have raised the question “What should we do to be able to calculate and report our emissions accurately?” and with this paper, we are trying to answer that question with our Software Company Scopes model.

Verifying the model needs academy-industry collaboration with both ICT companies and companies that calculate CO₂ emissions based on the GHG Protocol. Validating the model with a larger sample of companies can show its strengths and weaknesses and will open the way for future adjustments if needed. This can be achieved by calculating pilot companies’ emissions and comparing the results from the model against current emission calculations. To be reliably validated, there needs to be collaboration with companies that have not considered these issues widely before.

We acknowledge that the model needs validation through case studies where it is applied to software-producing companies. We also acknowledge that the sample of five companies represents SMEs, and the model might need adjustments in large companies.

The important question to research more is to find the largest emission sources and the low-hanging fruits. The largest sources for software company's emissions can vary between different kinds of software companies, depending on variables such as whether the company operates on a B2B or B2C model; the type of the software in question, such as SaaS, licensed software product, or tailored software; and architecture choices such as modular or client-server architecture. According to our research, the largest sources of emissions in software companies are located in Scope Three.

The model also needs to be customized for, e.g., consulting companies, digital marketing companies, and ICT hardware and infrastructure companies, which have their own characteristics. Consulting companies especially have quite a varying array of services provided, which raises the need for customization.

References

1. Global e-Sustainability Initiative: SMART 2020: Enabling the low carbon economy in the information age (2008). <https://gesi.org/research/download/7>. Accessed 10 Nov 2023
2. Global e-Sustainability Initiative: #SMARTer2030. ICT Solutions for 21st Century Challenges (2015). https://smarter2030.gesi.org/downloads/Full_report.pdf. Accessed 10 Nov 2023
3. Freitag, C., Berners-Lee, M., Widdicks, K., Knowles, B., Blair, G.S., Friday, A.: The real climate and transformative impact of ICT: a critique of estimates, trends, and regulations. *Patterns* 2(9), 100340 (2021)
4. GHG Protocol homepage. <https://ghgprotocol.org/>. Accessed 08 Sept 2023
5. Kochanowska, M., Gagliardi, W.R., with reference to Ball, J.: The double diamond model: in pursuit of simplicity and flexibility. In: Raposo, D., Neves, J., Silva, J. (eds.) *Perspectives on Design II*. Springer Series in Design and Innovation, vol. 16, pp. 19–32. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-79879-6_2
6. ISO 14064-1:2018. Greenhouse gases Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals. <https://www.iso.org/standard/66453.html>. Accessed 10 Nov 2023
7. The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard. <https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf>. Accessed 17 Sept 2023
8. The Greenhouse Gas Protocol: Corporate Value Chain (Scope 3) Accounting and Reporting Standard. https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporting-Standard_041613_2.pdf. Accessed 17 Sept 2023
9. Kern, E., et al.: Sustainable software products-Towards assessment criteria for resource and energy efficiency. *Futur. Gener. Comput. Syst.* **86**, 199–210 (2018)
10. Naumann, S., Kern, E., Dick, M.: Classifying green software engineering - the GREENSOFT model. *Softwaretechnik-Trends* **33**, 18–19 (2014)
11. Taina, J.: How green is your software? In: Tyrväinen, P., Jansen, S., Cusumano, M.A. (eds.) *ICSOB 2010*. LNBIIP, vol. 51, pp. 151–162. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-13633-7_13
12. Bourque, P., Fairley, R.E.: Guide to the Software Engineering Body of Knowledge - SWEBOK V3.0. (2014)
13. Shenoy, S.S., Eeratta, R.: Green software development model: an approach towards sustainable software development. In: 2011 Annual IEEE India Conference, pp. 1–6. IEEE (2011)
14. Becker, C., et al.: Requirements: the key to sustainability. *IEEE Softw.* **33**(1), 56–65 (2015)

15. Condori-Fernandez, N., Lago, P.: Characterizing the contribution of quality requirements to software sustainability. *J. Syst. Softw.* **137**, 289–305 (2018)
16. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. *J. Manage. Inf. Syst.* **24**(3), 45–77 (2008). <https://doi.org/10.2753/MIS0742-1222240302>
17. Jagroep, E., van der Werf, J.M., Brinkkemper, S., Blom, L., van Vliet, R.: Extending software architecture views with an energy consumption perspective. *Computing* **99**(6), 553–573 (2017)
18. Venters, C.C., et al.: Software sustainability: research and practice from a software architecture viewpoint. *J. Syst. Softw.* **138**, 174–188 (2018)
19. Pereira, R., et al.: Ranking programming languages by energy efficiency. *Sci. Comput. Program.* **205**, 102609 (2021)
20. Zwinkau, A.: Resource awareness for efficiency in high-level programming languages, Karlsruhe Institute of Technology, Technical report, Karlsruhe, Nr. 12 (2012)
21. Pereira, R., et al.: Energy efficiency across programming languages: how do energy, time, and memory relate? In: Proceedings of the 10th ACM SIGPLAN International Conference on Software Language Engineering, pp. 256–267 (2017)
22. Salonen, L.: Finding Ecologically Sustainable Design Principles for Mobile Devices – Towards a Heuristic Model from a List of Good Design Practices (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Understanding Cost Dynamics of Serverless Computing: An Empirical Study

Muhammad Hamza¹ , Muhammad Azeem Akbar¹ , and Rafael Capilla^{2,3} 

¹ Software Engineering Department, Lappeenranta-Lahti University of Technology,
15210 Lappeenranta, Finland

{muhammad.hamza, azeem.akbar}@lut.fi

² Rey Juan Carlos University, Móstoles, Spain
rafael.capilla@urjc.es

³ Lappeenranta-Lahti University of Technology, Lappeenranta, Finland

Abstract. The advent of serverless computing has revolutionized the landscape of cloud computing, offering a new paradigm that enables developers to focus solely on their applications rather than managing and provisioning the underlying infrastructure. These applications involve integrating individual functions into a cohesive workflow for complex tasks. The pay-per-use model and nontransparent reporting by cloud providers make it difficult to estimate serverless costs, impeding informed business decisions. Existing research studies on serverless computing focus on performance optimization and state management, both from empirical and technical perspectives. However, the state-of-the-art shows a lack of empirical investigations on the understanding of the cost dynamics of serverless computing over traditional cloud computing. Therefore, this study delves into how organizations anticipate the costs of adopting serverless. It also aims to comprehend workload suitability and identify best practices for cost optimization of serverless applications. To this end, we conducted a qualitative (interviews) study with 15 experts from 8 companies involved in the migration and development of serverless systems. The findings revealed that, while serverless computing is highly suitable for unpredictable workloads, it may not be cost-effective for certain high-scale applications. The study also introduces a taxonomy for comparing the cost of adopting serverless versus traditional cloud.

Keywords: Cost Dynamics · Serverless Computing · Empirical Investigation

1 Introduction

The advent of serverless computing has revolutionized the landscape of cloud computing, offering a new paradigm that enables developers to focus solely on their applications rather than managing and provisioning the underlying infrastructure [1]. Function-as-a-service (FaaS), an implementation serverless pattern, enables developers to create an application function in the cloud that automatically triggers in response to an event [1]. Companies employing the serverless model only pay for the resources consumed by the application compared to the traditional cloud, where a resource needs to be pre-reserved regardless of usage.

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBP 500, pp. 456–470, 2024.

https://doi.org/10.1007/978-3-031-53227-6_32

According to a survey conducted by Gartner Group, over 75% of organizations have either already adopted serverless computing or plan to do so within the next two years [2]. Moreover, the serverless market will substantially grow from \$3 billion in 2017 to an approximate value of \$22 billion by 2025 [3]. However, transitioning to a serverless computing model presents several challenges (e.g., legacy system integration, cold start, state management), and understanding the cost implications and identifying suitable workloads are crucial for effective adoption [4].

There has been significant recent research sought to address various aspects of serverless such as serverless architectural design [5], development features, technological aspects, and performance characteristics of serverless platforms [6], etc.. For instance, Lin et al. [7] extensively discuss a serverless architecture, proposed a formal construct for defining serverless application workflows, and introduced the Probability Refined Critical Path Greedy algorithm (PRCP) to optimize both performance and cost. Also, Wen et al. [8] conducted a systematic literature review and highlighted the benefits of serverless computing, its performance optimization, commonly used platforms, research trends, and promising opportunities in the field. However, to the best of our knowledge, no empirical study extensively investigated the systems transitioned to serverless computing or greenfield development. This includes aspects such as predicting serverless cost, serverless workload applicability, and cost optimization. Furthermore, there is a lack of taxonomy to compare the cost of adopting serverless and traditional cloud computing.

Therefore, this study investigates companies' decision-making process to determine the cost-effectiveness of adopting serverless computing. It also evaluates the suitability of various workloads for serverless computing. Additionally, the research identifies factors that contribute to high costs in serverless applications and explores the practices to optimize them. To this end, we analyzed eight systems that have successfully transitioned to serverless computing by conducting 15 interviews with industry professionals. In addition to our empirical analysis, we developed a taxonomy for comparing the cost of adopting serverless and traditional cloud computing.

Following, we presented three research questions that guided our study:

RQ1: How do companies estimate the cost of adopting serverless computing?

RQ2: Which specific types of workloads are best suited for serverless computing?

RQ3: What factors may increase the cost, and how can they be optimized?

The paper is structured as follows: Sect. 2 delves into related work, Sect. 3 outlines the research method, Sect. 4 discusses the results, Sect. 5 introduces the taxonomy on cost components, and Sect. 7 concludes the study.

2 Related Work

The existing studies have discussed different aspects of serverless computing, including architectural design, performance improvement, technological aspects, testing and debugging [9, 10], and empirical investigations [11–13].

Wen et al. [11] analyzed 619 discussions from the stack overflow repository. Their study uncovered the challenges (e.g., function configuration, package integration, function invocation) that developers face when developing a serverless application. Similarly,

Eskandani and Salvaneschi [12] provided insight into the FaaS ecosystem by analyzing the 2k real-world open-source applications developed using a serverless platform. The study collected open-source applications from GitHub and explores aspects like the growth rate of serverless architecture, architectural design, and common use cases. A similar study conducted by Esimann et al. [13] analyzed 16 characteristics that described why and when successful adopters are using serverless applications, and how they are building them by analyzing GitHub serverless projects [12].

Additionally, Adam et al. [14] propose guidelines for migrating to FaaS, aiming to optimize serverless functions to reduce memory consumption and running costs by conducting local experiments with their application. Another study conducted by Tarek et al. [15] developed an algorithm to optimize the cost of serverless applications through function fusion and placement. Similarly, Anil et al. [16] evaluated the AWS (Amazon Web Services) step function orchestrator concerning its performance and cost by conducting a series of experiments. Adzic and Chatley [17] conducted two industrial case studies from early adopters, demonstrating how transitioning an application to the Lambda deployment architecture reduced hosting costs. Their study did not present the cost optimization practices for companies.

Our study differs from the previous ones as we empirically investigate how organizations anticipate the cost implications of serverless computing. It also evaluates the suitability of various workloads for serverless computing. Additionally, the study identifies factors that contribute to high costs in serverless applications and explores the practices to optimize them. The existing studies did not cover these aspects of serverless computing.

3 Research Methodology

We employed a qualitative research method, specifically semi-structured interviews [18], to fulfill the objective of this study. Qualitative approaches aim to understand real-world situations, deal directly with complex issues, and are useful in answering “how” questions in the study [18]. The interviews were undertaken with 15 industrial participants who have experience in migrating legacy systems to serverless architectures or in developing serverless systems from scratch.

3.1 Data Collection

Interview Instruments. The semi-structured interview guide was developed based on the research questions following the guidelines of Robinson [19]. The interview guide covers demographic information, strategies followed by companies to understand the cost dynamics of serverless, serverless workload applicability, and strategies for optimizing application cost. The first and second authors were involved in developing the interview questions. The interview guide can be found at¹.

Participants Recruitment. The first two authors attended seven technology innovation industrial meetups where companies participated to share their success stories.

¹ <https://tinyurl.com/2kdraumf>.

Both authors randomly contacted industrial practitioners and asked them whether they employed serverless computing in their industry. In addition, the second author contacted the targeted population by leveraging social media platforms (e.g., LinkedIn, ResearchGate). A total of 38 participants were contacted, of which 15 were selected for the interview. We adopted a defined set of acceptance criteria for selecting our interviewees and case organizations. Mainly, our participants are (a) professional software engineers (b) who have participated in a serverless migration project within their professional scope or developed greenfield serverless application.

We finally shared the interview script with the practitioners beforehand to familiarize ourselves with the study. We interviewed 15 professionals from 4 countries (Finland, Netherlands, UAE, Pakistan) working at medium and large companies in different business domains. The first author conducted all the interviews online using Zoom and Microsoft Teams platforms. The interviews lasted for ~40 to ~55 min on average. The recorded interviews were transcribed for further analysis (Fig. 1).

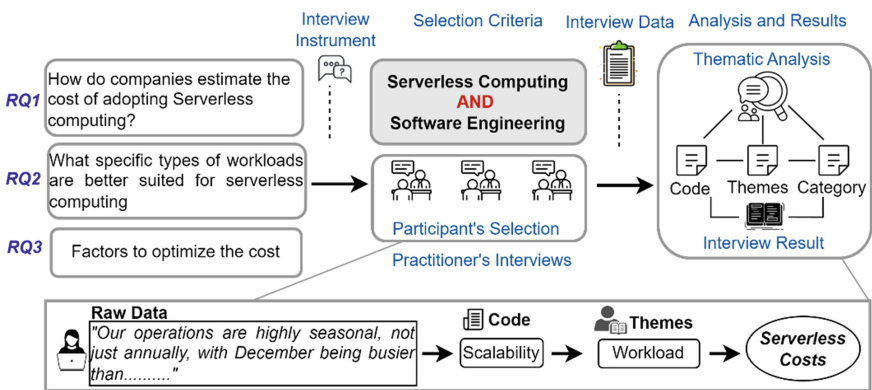


Fig. 1. Research Methodology

3.2 Data Analysis

This study used a thematic analysis approach to identify, analyze, and report the findings [20]. The thematic analysis enabled us to identify decision-making practices, workload applicability, and cost optimization practices, which were subsequently mapped into themes. We utilized NVivo² qualitative data analysis tool to identify and categorize the codes into themes. Initially, we meticulously read the interview transcriptions and made observational notes without establishing codes. After familiarization, we began coding the transcriptions, scrutinizing, and categorizing the resultant codes under the main themes. The main themes were decision, workload applicability, and cost optimization. The coding part was revisited repeatedly, and statements with similar meanings, but different phrasing were connected.

² <https://support.qsrinternational.com/s/>

Table 1. Company's demographics.

Company	ID	Domain	Employees
Co.1	S1	Logistics services	37365
Co.2	S2	E-commerce	15000
Co.3	S3	Web-applications	14500
Co.4	S4	E-commerce	9500
Co.5	S5	E-commerce	700
Co.6	S6	AI & Security Services	536
Co.7	S7	Smart mobility and security	20
Co.8	S8	E-commerce	3500

4 Results and Discussion

We conducted a comprehensive thematic analysis to obtain our results. Codes were extracted from interview transcripts and subsequently mapped into themes. These codes are denoted as C1, C2, C3, etc., while the corresponding themes are labeled T1, T2, and T3. Figure 2 provides a detailed representation of all identified codes and themes.

4.1 T1: Estimating Serverless Cost (RQ1)

In this section we present the practices practitioners employ to assess the cost of adopting serverless computing. Companies conduct a thorough cost analysis comparing the current infrastructure costs with the projected costs of serverless architecture. The following are the strategies reported by interviewed participants to predict the cost of serverless.

C1: Understanding Systems Nature. Serverless charges based on the pay-per-use model as compared to the traditional cloud. Therefore, understanding the nature and workload of the system is crucial before adopting a serverless model. The interviews revealed that serverless is the best fit for a system that receives a highly unpredictable workload. Many of the systems investigated follow an event-driven style. For instance, participant P1 stated, “*Our operations are highly seasonal, not just annually, with December being busier than June, but [...]. Given this variability, a serverless, event-driven architecture makes sense. It scales with the events, and we only pay for the events we use, reducing costs during off-peak times*”. In such scenarios, companies are compelled to over-provision each service, resulting in substantial resource wastage due to unused CPU utilization. Therefore, our interviewed participant assisted in assessing the workload of the system and monitoring the resource utilization of servers to decide to adopt serverless P1 further stated, “*It’s quite costly, and it genuinely pains me to witness an AWS account operating hundreds of EC2 instances, each running at less than 5% CPU utilization*”.

C2: Focusing on Unit Economics. Unit economics can guide the decision to adopt serverless models by comparing the cost per unit of request between current and serverless architectures. In this case, 8 out of 15 participants agreed that doing the unit analysis

can help make informed decisions for adopting serverless in terms of cost-effectiveness. If serverless offers a lower cost per unit, it may be a cost-effective choice P3 stated, *“I’ve realized the importance of understanding the unit economics of the systems we build. By identifying the cost per unit of value - for instance, the cost per scan in a security website scanning system - we can better manage resources and demonstrate our true profitability. This approach is particularly beneficial in serverless architectures”*. Another participant P8 stated that *“Based on my calculations, handling 100 million requests via API Gateway and Lambda is cost-effective and more scalable compared to traditional clusters.”*

C3: Testing Costly Components with Serverless. Participants identified the most expensive components in a large monolithic system and employed domain-driven design to extract these components. They migrated these isolated components to a serverless architecture to assess whether this transition is cost-effective. For instance, P9 stated, *“We advocate for serverless rightsizing. We start by identifying the most expensive components in a legacy system and strategically migrating them to a serverless architecture. An automated cost-benefit analysis accompanies this process, providing solid justification for the transition. In our experience with serverless, we’ve seen the potential for substantial returns, even up to a 100-fold return on investment”*. Therefore, testing the costly component with serverless and gradually migrating is the best practice reported by the participants to be cost-effective.

C4: Enabling a Cost-Conscious Team. Empowering a cost-conscious team is a crucial step in evaluating the cost implications of adopting serverless architecture and making an informed decision about the serverless in terms of cost-effectiveness. As stated by P13: *“So you know, you need someone who understands both the finance side of things, as well as the technical side of things to really sort of kind of appreciate some of the total cost of ownership applications that serverless has”*.

C5: Serverless First Mindset. Organizations developing greenfield projects must go with a serverless first mindset P15 stated: *“I think if you’re a startup and you’re building on AWS, it just doesn’t make sense for you to do anything than serverless [...] You know, the cost of containers is so much more operations work, and probably must hire some specialists, just to look after your container environment”*. However, applications having high throughput could not be cost-effective in serverless computing as stated by P14 *“The funny thing is that a lot of the enterprises, they don’t really have that high throughput applications where you will be significantly more expensive to run on serverless compared to containers”*. However, to effectively understand the cost-effectiveness of serverless computing, it’s crucial to deeply understand the nature of the system, emphasizing on unit economics, assessing the costly components of legacy application, and testing with serverless, and cultivating a team that is acutely cost-aware.

4.2 Interview Cases Description (RQ2)

This section delves into the case studies of systems that have either migrated to a serverless architecture or were developed greenfield serverless systems. We investigated eight systems by interviewing 15 participants, which we refer to as ‘S1-S8,’ from companies labeled as ‘Co.1–Co.8’ (where ‘Co’ stands for ‘Company’ and ‘S’ stands for ‘System’). The details of participating companies (Co.1–Co.8) of different sizes and domains are

shown in Table 1 and Table 2. We presented a short introduction to each system naming them S1–S8 from companies Co.1–Co.8. Furthermore, we understand the type of traffic the systems were receiving (e.g., unpredictable, or spiky traffic, predictable traffic). We derived three codes (C6: unpredictable or spiky workload, C7: workload having less than 1000 req/s, C8: predictable workload) by analyzing the eight systems and mapped into themes T2: workload applicability presented in Fig. 2.

Table 2. Participant’s demographic

Participants	Participant’s Role	Professional Experience	Serverless Experience
P1	Architect	18	5
P2	Architect	16	5
P3	Architect	9	3
P4	Architect	13	5
P5	Developer	8	3
P6	Architect	15	4
P7	Lead Engineer	5	2
P8	Software Engineer	5	2
P9	Team Manager	5	2
P10	Software Engineer	6	2
P11	Architect	18	5
P12	Architect	16	5
P13	Architect	9	4
P14	Architect	13	5
P15	Developer	8	4

Co.1-S1 Logistic Management System. Co.1 is a large-scale enterprise offering logistics services, including domestic and international mail and parcel delivery and e-commerce solutions. The system was facing seasonal traffic, causing the organization to handle the underlying operational overhead. P1 stated that: *“Our operations are highly seasonal, not just annually, with December being busier than June, but also weekly and daily. For instance, Tuesdays are busier than Mondays, and there’s a surge of traffic around 4:00 p.m. and 5:00 p.m. A serverless architecture scales with events and cuts costs during off-peak times, [...]”*. This company first evaluated the system’s nature and then conducted a proof-of-concept (POC). Additionally, they identified the expensive components in a traditional cloud setting and tested them with a serverless approach. The company was able to cut costs by 80% and reduce delivery times from months to minutes for its e-commerce API services migrating to serverless P1 stated, *“The business case became evident when we realized that by transitioning from a fixed instance and discarding our old data-management software, we could reduce our data-management platform costs by at least 80%”*.

Co.2-S2 E-Commerce. The company simplifies daily life for thousands of satisfied customers by offering a wide range of products for everyday needs and special occasions. They offer delivery at a time that suits the customer, often on the same day. According to P2: *“So we have very low traffic at night, steady traffic during the day, small spike at lunch, goes up in the evening, and then it dies off at midnight.”* So, the system faced seasonal traffic in peak times and was facing challenges managing servers. They extracted components from the legacy application and tested with serverless. They did the unit calculation of the received traffic and decided serverless could reduce the cost and improve the scalability. The migration reduces significant costs and operation overhead.

Co.3-S3 Digital Product Development. The company offers a variety of digital services designed to help businesses thrive in a digital-centric landscape using their web-based platform. The company has predictable traffic, handling millions of requests per month and wanted to reduce the operational overheads. They leveraged the serverless and reduced the cost from 1 thousand dollars to five hundred as stated by P7: *“By migrating from EC2 to serverless, we drastically reduced our costs while still providing the same services”*.

Co.4-S4 Pitch Decker. This company helps startups with various aspects, such as pitching to investors and getting up and running. Initially, they used AWS EC2 instances for hosting but encountered scalability and maintenance issues. P4 stated: *“We struggled with determining when to scale up or down as our app, not being time-sensitive or event-driven, didn’t present predictable traffic spikes [...]”*. They were spending a lot more time managing the underlying infrastructure rather than focusing on the business logic. Therefore, migrating to serverless reduced the operational overhead as the company does not want to hire a DevOps team.

Co.5-S5 E-commerce. The company specializes in providing custom apparel and accessories to its customers using its design tools. The company was facing the high cost of managing the servers and scalability issues as they received unpredictable workloads during the seasonal time stated by P5: *“We had to move that to a sort of more performance, more scalable system, where we didn’t have to sort of keep scaling up these EC2 instances”*. They moved a key part of their design architecture from an app to a Node-based Lambda. This transition resulted in 90% cost savings and improved performance and scalability. *“We got like immediate cost savings as well as sort of a capability expansion”*.

Co.6-S6 AI Virtual Assistant. The company provides financial services with artificial intelligence and machine learning (AI/ML) solutions. The system can read, comprehend, and draw conclusions based on context to mimic cognitive thinking and build expertise over time. Their previous infrastructure (EC2) was becoming increasingly expensive, with their monthly cloud bill rising. The system consistently manages a steady and predictable volume of traffic. However, migrating to serverless reduced the cost significantly, as stated by P12: *“After assessing the serverless pay-per-use model, we opted to implement it, resulting in an impressive cost reduction of approximately 87%”*.

Co.7-S7 Smart Mobility System. The startup company developed a smart mobility data generation system. This system involves collecting data from mobile phones and sending it to the startup’s backend infrastructure. The startup wants to develop a system

where they reduce the cost of the system and does not manage underlying infrastructure, as stated by P10: *“The need for scalability and flexibility in their operations was paramount. We want to get rid of like the time we spent on managing servers”*. The company evaluated that the nature of its system is event-driven and will grow exponentially, so it decided to go with a serverless first mindset.

Co.8–S8 E-Commerce. The company provides e-commerce services mainly for ordering food and grocery items. Initially, the company had a big monolithic system and faced issues such as scalability during peak seasons as their traffic was unpredictable, faster time to market, and high operation overhead (e.g., managing EC2 instances). These issues led to increased costs. The P11 stated that *“We wanted to create something we could own and rapidly iterate on. However, I was concerned about scaling and didn’t want to deal with potential EC2 server crashes or backend container issues”*. However, migration to serverless improved the scalability and reduced operational overhead and overall cost significantly.

Most of the interview systems (5) and participants (11) reported that migrating the unpredictable or spiky workload to serverless would significantly reduce the cost. However, three systems had a predictable workload and stated that they reduced the cost of going serverless P9: *“While running containers might seem cheaper initially, the hidden costs of expertise, maintenance, and scalability can quickly add up. Serverless, despite a potentially higher bill, can save costs by eliminating the need for specialized skills and infrastructure management”*. So, there is a tradeoff going serverless. Six out of 15 participants agreed that there are no universal solutions, only tradeoffs, and the choice between serverless and containers depends on the specific context and requirements. While serverless theoretically offers infinite scalability, it has a burst concurrency limit stated by P13 *“you know at high scale (1000 + req/s), services like API Gateway and Lambda can be more expensive than running containers on ECS. Lambda may also not be suitable for long-running tasks that take more than 15 min or applications with strict latency requirements”*, making it unsuitable for certain stabilized high-scale applications.

4.3 Cost Optimization Practices (RQ3)

This section highlights the primary factors increasing the costs in serverless architecture and outlines some solutions to optimize these costs from the practitioner’s perspective.

C9: Recursive Function Calling. Refers to the situation where a serverless function triggers itself, directly or indirectly, causing a loop of invocations. This recursive triggering can result in many function invocations, increasing the overall computation time and potentially leading to unexpectedly high costs. Practitioner P 6 stated: *“During our work with a customer’s system migration, an unexpected cost spike occurred due to code calling the KMS API millions of times, which they were unaware of until we generated an alert”*. However, practitioners employ different practices, including error handling and retry policies, use of idempotency keys, circuit breaker pattern, rate limiting, and recursive loop detection to handle the recursive function calling.

C10: Unused Functions. Functions deployed but not invoked or used over a significant period occupy resources and may incur costs even if they are not actively serving requests. According to P8: *“We periodically review and delete unused Lambda functions*

and associated resources (e.g., API Gateway, DynamoDB tables, S3 buckets) to minimize unnecessary costs”.

C11: Unintended Logging. Refers to excessive log data generation due to debug-level logging, verbose logging, or configuration mistakes. This not only incurs unnecessary costs for data storage and transfer in services but also complicates the process of extracting useful information from the logs. “*We experienced excessive data collection in monitoring solutions like Datadog that lead to significant costs, especially as usage scales from development to production [...]*”.

C12: Inefficient Data Access Patterns. This leads to a situation where developers might store a relatively small amount of data external database, but they’re accessing or retrieving that data frequently. If the data is being retrieved millions of times a day, even if it’s a small amount, the costs for these API requests can add up quickly and become significant. P11 stated: “*Inefficient access patterns in S3, such as frequent API calls to retrieve small amounts of data, can significantly increase costs, even if the stored data volume is low*”. Our interviewed practitioners mitigate this problem by considering data access patterns and optimizing them to minimize the number of API requests. This might involve using caching, batch retrieval of data, or redesigning their application to reduce the frequency of data retrieval.

C13: Denial of Wallet Attack. In this attack, an attacker intentionally triggers many function executions in a serverless application to inflate the application’s operational costs. According to P9: “*We’re aware of the risk of Denial-of-Wallet attacks in serverless architectures. Rapid scaling can lead to significant costs, so we ensure to have alerts and alarms in place to prevent unexpected expenses*”.

Our interview revealed the practices that need to be adopted to optimize the cost of serverless applications.

C14: Function Right Sizing. Involves matching the allocated resources to the actual usage of your functions. Over-provisioning can lead to unnecessary costs, while under-provisioning can hurt performance, as stated by “*We’ve learned that finding the ‘right sizing’ for Lambda functions is crucial - balancing performance and cost by continuously fine-tuning settings like memory allocation*”.

C15: Provisioned Concurrency. Keeps functions initialized and ready to respond instantly for reserved instances. However, mismatching the reserved instances can lead to high cost. “*We prioritize optimizing cost and performance in operations [...] understanding concurrency patterns and behavior is essential for effective implementation*”.

C16: Observing System Metrics. System metrics can provide insights into the application’s performance and resource usage. This information can guide optimization efforts and help identify potential cost savings. According to P9: “*You just keep an eye on things, make sure that you haven’t missed any alerts or stuff like that, which is great when you’ve got talking about the system of time and it being an operational thing for cost data because there’s such a big delay*”.

C17: Direct Integration. Involves connecting services directly instead of using intermediary services. This can reduce latency, improve performance, and lower costs “*I have personally witnessed the advantages of directly integrating serverless services, which can effectively decrease Lambda costs*”.

C18: Avoiding Idle Time. Refers to the period when resources are allocated but not actively used. In a serverless architecture, you’re billed for the computing time you consume, so reducing idle time can significantly cut costs. “We know it’s vital to avoid idle wait time in Lambda functions; using Lambda as an orchestrator for long gaps incurs unnecessary costs, so we optimize by focusing on active processing tasks”.

Apart from these, practitioners also highlighted that optimizing the code of the function, enabling billing alerts, giving developers billing access, and evaluating third-party tooling can significantly improve the optimizations and cost of the serverless application.

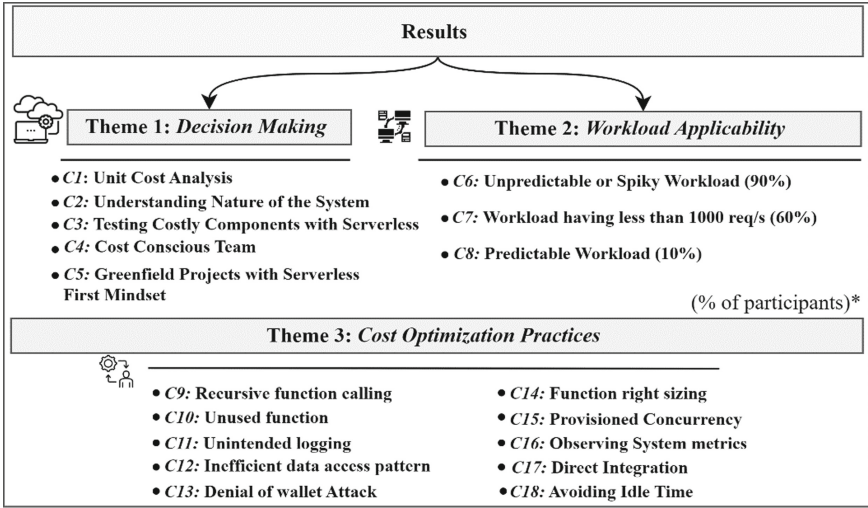


Fig. 2. Results from thematic analysis

5 Taxonomy of Factors Comparing the Cost of Ownership

In this section, we presented a taxonomy of factors comparing the cost of ownership between serverless and traditional cloud computing. The model is mainly divided into three components (i.e., infrastructure, development, and maintenance). We explained these components in detail and compared them with serverless and traditional cloud. This comparative analysis aims to provide organizations with insights to make informed decisions, comparing their cost of ownership in either computing model.

Infrastructure Cost. Incurred when utilizing a cloud service provider for hosting an application workload. The infrastructure cost comprises the computing, storage, and network services the host application consumes. On the traditional cloud, the computing cost is calculated based on the reserved instances for a specific period, whereas in serverless computing, the cost is calculated by actual execution time, achieving the 100% utilization of the resources. Our empirical analysis showed that systems on EC2 instances or servers were not fully utilizing their computational resources leading to waste of resources and operational overhead. Furthermore, utilizing services such as

load balancing, fault tolerance, and security cost extra charges on the traditional clou, whereas serverless architecturally provides these services. Organizations further need to evaluate the cost of database (e.g., compare the cost of querying NoSQL, such as MongoDB and DynamoDB). Therefore, organizations need to compare the computing, storage, and network cost of serverless and traditional cloud to make an informed decision (Fig. 3).

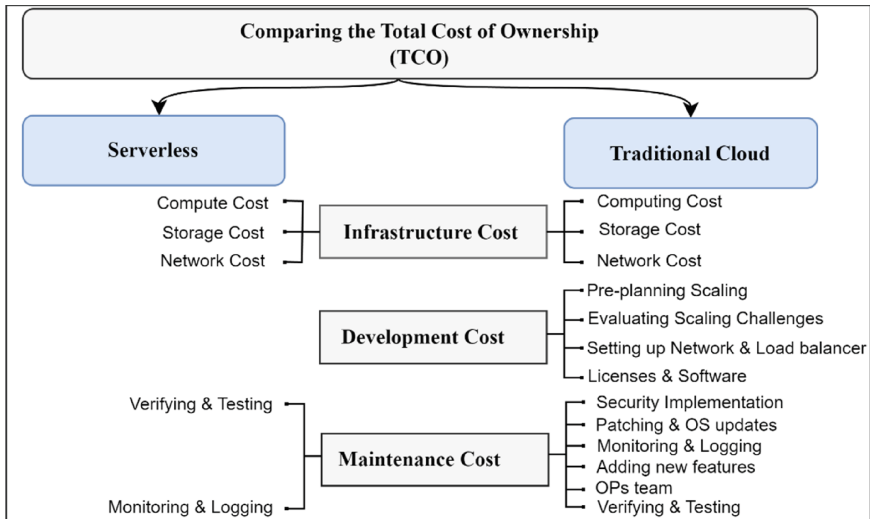


Fig. 3. A taxonomy of factors influencing the cost

Development Cost. This refers to the effort and time spent designing and developing applications on cloud-based services. In traditional cloud, developers need to evaluate how the architecture would scale over time. The developer must focus on utilizing the resources in scaling up and down in a traditional cloud environment. Developers utilizing EC2 instances are required to dedicate significant time to assess potential scalability challenges within the IT architecture and decide on necessary tradeoffs in the preliminary stages. This incurred the cost of planning the resources and time. In addition, the developer must spend more time setting up a network, load balancer, purchasing licenses and software, and planning availability. In contrast, serverless computing leads developers to build the application without worrying about planning scaling and the deployment of the application. The cost of planning has become negligible in serverless computing.

Maintenance Cost. This pertains to the ongoing cost required for running and maintaining an application. In the serverless, developers or operation teams do not need to maintain the application (e.g., patching and operating system updates). However, applications developed using cloud containers require extra work and labor to handle the application (i.e., DevOps team). The maintenance and operational costs become negligible in serverless computing compared to traditional cloud servers. Thus, leading to significantly lower costs overall and reducing the scalability issues and operational overhead.

Organizations considering adopting serverless or traditional cloud need to evaluate each component to make informed decisions.

6 Threat to Validity

Several potential threats could impact the validity of the results of this study. These threats are typically categorized into four primary categories: internal validity, construct validity, external validity, and conclusion validity [21].

Internal Validity: Refers to the degree to which specific factors influence methodological robustness. The first threat to this study is the participants' understanding of the interview questions. To mitigate this threat, we conducted pilot interviews with professionals from our network and provided them interview questions in advance. This ensured that the questions were both understandable and readable. We revised the interview questions based on the participants' feedback. The final interview preamble is provided in this study.

Construct Validity: Refers to the degree to which the research constructs are adequately substantiated and interpreted. The core constructs are the interview participants' viewpoints on the migration or adoption of serverless technology in the context of cost. The verifiability of the construct is considered the limitation of thematic analysis. Therefore, we followed a rigorous and step-by-step research method process and gave examples in quotations from the collected data (e.g., interviews).

External Validity: Refers to the generalizability of the results. The sample size and sampling approach of this study may not generalize the findings. A common threat can arise that serverless is not widely adopted in the industry. Similarly, migration to serverless is not well established in the practice. Finding the potential sample size was challenging for us. We mitigated this threat by using possible sources such as social media platforms (e.g., LinkedIn, ResearchGate) and attending seven industrial meetups to find the potential population. We collected data from 4 countries across two continents from participants with diverse experience in various industrial domains and in serverless.

Conclusion Validity: Refers to the factors that impact the trustworthiness of the study conclusion. To mitigate this threat, we conducted weekly meetings to develop the interview instruments and data analysis process. We reviewed the data based on the weekly discussion to improve the analysis process. Finally, we conducted a brainstorming session to draw the findings and conclusion of this study.

7 Conclusion and Future Work

Serverless computing presents a promising avenue for organizations to optimize costs and improve efficiency by minimizing scalability issues and operational overhead. However, successfully transitioning to serverless computing requires a deep understanding of cost implications and workload suitability. To this end, our study comprehensively analyzes cost optimization and workload suitability in serverless computing. Through an

empirical investigation of eight systems and 15 interviews with industry professionals, we identified how companies predict the cost of adopting serverless, workload suitability, and factors that affect the cost of serverless applications. Furthermore, we presented a theoretical model for understanding the cost of serverless compared with traditional cloud.

Our study revealed that most of the organizations do unit cost economics and migrating legacy components to serverless to understand the cost benefits of serverless. Moreover, most of the systems and interviewers stated that serverless is suitable for highly predictable workload, where developers need to spend most of the time provisioning the underlying infrastructure. Three interviews stated that, while serverless theoretically offers infinite scalability, it has a burst concurrency limit that could not be cost-effective for certain stabilized high-scale applications. However, all the suggested developing greenfield projects with the serverless first mindset. Further they assisted transitioning to containers when it becomes more cost-effective. In addition, this study also identified factors that can increase the cost and strategies used to optimize the application cost. Finally, we developed a taxonomy for evaluating the cost of serverless versus traditional cloud computing. This taxonomy serves as a valuable tool for organizations, helping them make more informed decisions about which cloud computing model is most cost-effective for their specific needs.

As future work, we plan to extend our findings by mining Q&A repositories and conducting a survey with a larger number of industrial practitioners. Further, we aim to develop a comprehensive theory that explains how decisions are made at every stage of migrating to serverless computing—from planning and development to deployment.

References

1. Jonas, E., et al.: Cloud programming simplified: a berkeley view on serverless computing, arXiv preprint [arXiv:1902.03383](https://arxiv.org/abs/1902.03383) (2019)
2. Raj, et al.: Gartner Magic Quadrant for Cloud Infrastructure and Platform Services. <https://tinyurl.com/37a8pznt>
3. Datadog: The state of serverless (2022). <https://www.datadoghq.com/state-of-serverless/>. Accessed 12 July 2023
4. Ko, H., Pack, S., Leung, V.C.: Performance optimization of serverless computing for latency-guaranteed and energy-efficient task offloading in energy harvesting industrial IoT. *IEEE Internet Things J.* **10**, 1897–1907 (2021)
5. Hamza, M.: Software architecture design of a serverless system. In: Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering, pp. 304–306 (2023)
6. Yussupov, V., Breitenbücher, U., Leymann, F., Wurster, M.: A systematic mapping study on engineering function-as-a-service platforms and tools. In: Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing, pp. 229–240 (2019)
7. Lin, C., Khazaei, H.: Modeling and optimization of performance and cost of serverless applications. *IEEE Trans. Parallel Distrib. Syst.* **32**(3), 615–632 (2020)
8. Wen, J., Chen, Z., Jin, X., Liu, X.: Rise of the planet of serverless computing: a systematic review. *ACM Trans. Softw. Eng. Methodol.* **32**(5), 1–61 (2023). <https://doi.org/10.1145/3579643>

9. Lenarduzzi, V., Panichella, A.: Serverless testing: tool vendors' and experts' points of view. *IEEE Softw.* **38**(1), 54–60 (2020)
10. Rinta-Jaskari, E., Allen, C., Meghla, T., Taibi, D.: Testing approaches and tools for AWS lambda serverless-based applications. In: *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pp. 686–692. IEEE (2022)
11. Wen, J., et al.: An empirical study on challenges of application development in serverless computing. In: *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pp. 416–428 (2021)
12. Eskandani, N., Salvaneschi, G.: The uphill journey of FaaS in the open-source community. *J. Syst. Softw.* **198**, 111589 (2023)
13. Eismann, S., et al.: The state of serverless applications: collection, characterization, and community consensus. *IEEE Trans. Software Eng.* **48**(10), 4152–4166 (2021)
14. Stafford, A., Toosi, F.G., Mjeda, A.: Cost-aware migration to functions-as-a-service architecture. In: *ECSA (Companion)* (2021)
15. Elgamal, T.: Costless: optimizing cost of serverless computing through function fusion and placement. In: *2018 IEEE/ACM Symposium on Edge Computing (SEC)*, pp. 300–312. IEEE (2018)
16. Mathew, A., Andrikopoulos, V., Blaauw, F.J.: Exploring the cost and performance benefits of AWS step functions using a data processing pipeline. In: *Proceedings of the 14th IEEE/ACM International Conference on Utility and Cloud Computing*, pp. 1–10 (2021)
17. Adzic, G., Chatley, R.: Serverless computing: economic and architectural impact. In: *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, pp. 884–889 (2017)
18. Benbasat, I., Goldstein, D.K., Mead, M.: The case research strategy in studies of information systems. *MIS Q.* **11**(3), 369 (1987). <https://doi.org/10.2307/248684>
19. Robinson, O.C.: Sampling in interview-based qualitative research: a theoretical and practical guide. *Qual. Res. Psychol.* **11**(1), 25–41 (2014)
20. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qual. Res. Psychol.* **3**(2), 77–101 (2006)
21. Zhou, X., Jin, Y., Zhang, H., Li, S., Huang, X.: A map of threats to validity of systematic literature reviews in software engineering. In: *2016 23rd Asia-Pacific Software Engineering Conference (APSEC)*, pp. 153–160. IEEE (2016)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Quantum Software Ecosystem: Stakeholders, Interactions and Challenges

Vlad Stirbu^(✉)  and Tommi Mikkonen 

University of Jyväskylä, Jyväskylä, Finland
{vlad.a.stirbu,tommi.j.mikkonen}@jyu.fi

Abstract. The emergence of quantum computing proposes a revolutionary paradigm that can radically transform numerous scientific and industrial application domains. The ability of quantum computers to scale computations imply better performance and efficiency for certain algorithmic tasks than current computers provide. However, to gain benefit from such improvement, quantum computers must be integrated with existing software systems, a process that is not straightforward. In this paper, we investigate the quantum computing ecosystem and the stakeholders involved in building larger hybrid classical-quantum systems. In addition, we discuss the challenges that are emerging at the horizon as the field of quantum computing becomes more mature.

Keywords: Quantum software · Quantum ecosystem · Value chain

1 Introduction

Quantum computing holds great promise as a revolutionary technology that has the potential to transform various fields. By harnessing the principles of quantum mechanics, quantum computers can perform complex calculations and solve problems that are currently intractable for classical computers. This promises breakthroughs in areas such as cryptography, optimization, drug discovery, materials science, and machine learning. Quantum computing's ability to leverage quantum mechanics properties like superposition, interference and entanglement can unlock significant speedups and enable more accurate simulations of quantum systems.

The development of quantum software faces numerous challenges that need to be addressed for harnessing the power of quantum computing effectively. Firstly, the limited availability and instability of quantum hardware pose significant obstacles. Quantum computers are prone to errors and noise, necessitating the development of robust error correction techniques. Further, quantum programming languages and tools are still in their nascent stages, requiring improvements to facilitate efficient software development. More, the scarcity of skilled quantum

This work has been supported by the Academy of Finland (project DEQSE 349945) and Business Finland (project TORQS 8582/31/2022).

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBP 500, pp. 471–477, 2024.

https://doi.org/10.1007/978-3-031-53227-6_33

software developers and a lack of standardization hinder the widespread adoption of quantum software. As quantum systems scale, the complexity of designing and optimizing quantum algorithms increases, demanding novel approaches to algorithm design and optimization. Addressing these challenges is crucial for realizing the full potential of quantum computing and enabling the development of practical quantum software applications.

In this paper, we delve into the realm of the quantum software ecosystem and examine the interconnections among its stakeholders. Our focus centers on the intricate interplay between these entities, and we pinpoint their areas of influence within the technology stack. Ultimately, our objective is to provide both established stakeholders and emerging participants with insights that can inform their strategic decision-making.

The rest of the paper is structured as follows. The background is provided in Sect. 2. The ecosystem overview is presented in Sect. 3. The discussion of the value stream within the ecosystem is provided in Sect. 4. Concluding remarks are provided in Sect. 5.

2 Background

2.1 Qubit Implementation

The current candidates for building general-purpose quantum computers, as listed in Table 1, fall under the category of Noisy Intermediate-Scale Quantum (NISQ) systems. Although these quantum computers are not yet advanced enough to achieve fault-tolerance or reach the scale required for quantum supremacy, they provide an experimentation platform to develop new generations of hardware and quantum algorithms and validate quantum technology in real world use cases. Whether a quantum computer is general-purpose or specialized, the selection of quantum qubit implementation technology can significantly enhance hardware efficiency for specific problem classes. To make effective use of the hardware, application developers must consider these differences when designing and optimizing the software's functionality and operations.

2.2 Quantum Algorithms

Quantum algorithms are computational techniques specifically designed to harness the unique properties of quantum systems [2]. They offer significant advantages over classical algorithms in certain computational tasks. One key advantage is the ability to solve complex problems faster. For example, Shor's algorithm enables efficient factoring of large numbers, posing a potential threat to current encryption methods. Also, Grover's algorithm provides substantial speedup in searching large databases. Moreover, quantum algorithms can address optimization problems more effectively, leading to improved solutions in areas like portfolio optimization, logistics, and drug discovery, to name some concrete examples.

Table 1. Qubit implementation technologies.

Qubit Technology	Description	Applicability
Superconducting	Tiny superconducting materials are cooled to extremely low temperatures to manifest their quantum properties	General-purpose quantum computing, suitable for various types of problems
Trapped Ion	Ions are trapped within electromagnetic fields	General-purpose quantum computing, with potential for high coherence and low error rates
Photonic	Quantum information stored in photons can be manipulated and transmitted over long distances	General-purpose quantum computing, suitable for communication and cryptography applications
Annealing	Special purpose quantum computers designed to solve optimization problems	Specialized quantum computing, targeted at optimization and sampling problems
Topological	A new approach to quantum computing that leverages the properties of topological states of matter to create qubits. Topological qubits are based on collective properties of an ensemble of particles	General-purpose quantum computing, aimed at achieving fault-tolerant operations

2.3 Software

A typical quantum program performs a specialized task as part of a larger classical program. The quantum program is submitted as a batch task to a classical computer that controls the operation of the quantum computer. The classical computer schedules the task execution and provides the result to the classical program when the job completes. To support this process, numerous alternatives for tooling exist.

An application developer use tools like Qiskit¹ and Cirq² for writing, manipulating and optimizing quantum circuits. These Python libraries allow researchers and application developers to interact with nowadays' NISQ computers, allowing them to run quantum programs on a variety of simulators and hardware designs, abstracting away the complexities of low-level operations and allowing researchers and developers to focus on algorithm design and optimization.

Tools like TensorFlow Quantum³ and PennyLane⁴ play a crucial role in facilitating the development of machine learning quantum software. These frameworks provide the high-level abstractions and interfaces that bridge the gap between quantum computing and classical machine learning. They allow researchers and developers to integrate quantum algorithms seamlessly into machine learning development process by providing access to quantum simulators and hardware,

¹ <https://qiskit.org>.

² <https://quantumai.google/cirq>.

³ <https://www.tensorflow.org/quantum>.

⁴ <https://pennylane.ai>.

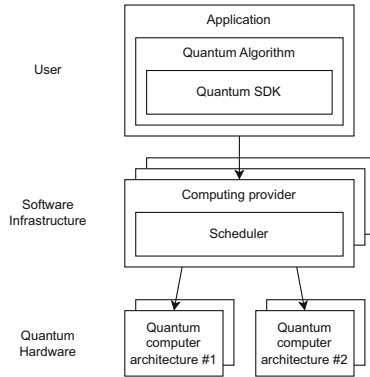


Fig. 1. Quantum stack layers and components.

as well as offering a range of quantum-friendly classical optimization techniques. TensorFlow Quantum leverages the power of Google’s TensorFlow ecosystem, enabling the combination of classical and quantum neural networks for hybrid quantum-classical machine learning models. PennyLane offers a unified framework for developing quantum machine learning algorithms, supporting various quantum devices and seamlessly integrating them with classical machine learning libraries.

Traditional cloud computing providers, such as AWS Bracket⁵, Azure Quantum⁶, Google Quantum AI⁷ or IBM Quantum⁸, offer comprehensive quantum development services. These services are designed to optimize the development process, with integrated tools like Jupyter⁹ notebooks and task schedulers. Developers can create quantum applications and algorithms across multiple hardware platforms simultaneously. This approach ensures flexibility, allowing fine-tune algorithms for specific systems while maintaining the ability to develop applications that are compatible with various quantum hardware platforms.

3 Ecosystem Layers and Stakeholders

The quantum ecosystem can be segmented into distinct functional layers, as illustrated in Fig. 1. The first one is the *user* layer, encompassing applications and supplementary software components crafted by third-party developers. This includes quantum algorithms and software development kits (SDKs) for quantum circuits, such as Cirq and Qiskit. The *infrastructure* layer, in contrast, comprises the software employed by computing providers to manage and execute quantum

⁵ <https://aws.amazon.com/braket/>.

⁶ <https://learn.microsoft.com/en-us/azure/quantum/>.

⁷ <https://quantumai.google>.

⁸ <https://quantum-computing.ibm.com>.

⁹ <https://jupyter.org>.

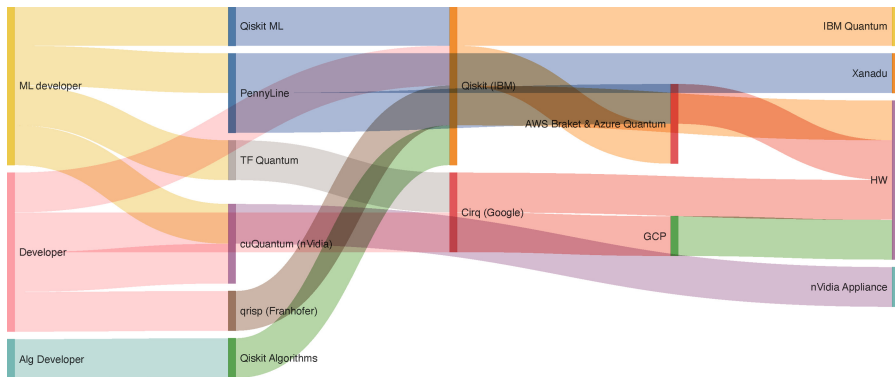


Fig. 2. Quantum ecosystem: stakeholders, software tools and interactions

computing tasks specified within the user layer. Finally, the *hardware* layer pertains to the physical hardware and accompanying control software essential for implementing the qubits required to execute quantum circuits.

From a stakeholder perspective, each functional layer is characterised by specific entities of interest. The user layer is primarily populated by the business and scientific stakeholders that commission the development of the respective applications. Typically, these applications use third-party algorithm libraries and quantum circuit SDKs. Quantum algorithm developers and researchers often contribute to these libraries as a means to disseminate their work. Similarly, the quantum circuit SDKs provide unique idioms to program quantum circuits making easy for developers to define and control the individual quantum gates. At the infrastructure layer, we find the major cloud computing providers and to a lesser extent the quantum hardware manufacturers. The hardware layer consists of the quantum computer manufacturers and the myriad of suppliers that provide the components for the respective hardware.

4 Discussion

Today, Cirq and Qiskit have established market dominance in the general purpose quantum computing. Similarly, PennyLane is the dominant ML specialized framework, besides Cirq and Qiskit. These frameworks provide strong control points for Google, IBM, and Xanadu, respectively, to control the programming space, see Fig. 2. Independent hardware manufacturers have to provide back-end implementations for these SDKs in order to enable application developers to write programs that use their devices. Similarly, frameworks like qrisp [3], which provides an alternative quantum circuit programming model, have to fold into the realities of the ecosystem and provide Qiskit-compatible back-end wrappers to be able to execute on existing quantum hardware.

As the race towards quantum supremacy is still in its infancy, the quantum hardware needs to evolve from the current computers that offer tens of qubits to

at least hundreds and being able to execute circuits with thousands of gates [1]. As the hardware development is resource intensive, the manufacturers might find themselves isolated into the lower layer of the stack, limited to providers of back-end implementations for the established programming frameworks. However, to be able to interact with developers they have to expose additional functionality at the appropriate layer in the upper software stack, above Qiskit or PennyLane for example.

The quantum computing community, deeply rooted in scientific principles, embraces collaboration and often adopts an open-source approach for many frameworks and software tools. Nevertheless, these projects are controlled by commercial interests, and open governance is often lacking or limited. A notable exception is QIR Alliance¹⁰, a Linux Foundation led effort aiming to develop standards for interoperability in the quantum compiler space. An area of special interest is tooling related to scheduling and execution, where the cloud providers have a clear advantage. An open source execution environment developed using an open governance model, similar to Kubernetes, would allow smaller players to operate quantum computing services in a cost efficient matter.

5 Conclusions

The emergence of quantum computing is spurring a new ecosystem, where quantum computers must be integrated with existing software systems and their development. In this paper, building on early research results and practical observations, we have mapped out the stakeholders and shed light on the dynamics within today's quantum software ecosystem. However, more in-depth investigation is needed for the exploration of stakeholders' unique interests and fundamental characteristics of the systems they provide and propose. To this end, our analysis of the quantum ecosystem, its stakeholders, and their interactions serves as a valuable starting point, setting the stage for deeper exploration and enhanced understanding of the quantum computing field.

References

1. Dalzell, A.M., Harrow, A.W., Koh, D.E., La Placa, R.L.: How many qubits are needed for quantum computational supremacy? *Quantum* **4**(264) 10 May 2020 <https://doi.org/10.22331/q-2020-05-11-264>
2. Montanaro, A.: Quantum algorithms: an overview. *NPJ Quant. Inform.* **2**(1), 15023 (Jan 2016). <https://doi.org/10.1038/npjqi.2015.23>
3. Seidel, R., Tcholchev, N., Bock, S., Hauswirth, M.: Uncomputation in the qrisp high-level quantum programming framework. In: Kutrib, M., Meyer, U. (eds.) *Reversible Computation*, pp. 150–165. Springer Nature Switzerland, Cham (2023)

¹⁰ <https://www.qir-alliance.org/alliance/>.






Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Dynamic Capabilities for Sustainable Digital Transformation Amid Crisis: Insights from Law Firms in Emerging Economy

Mikhail O. Adisa¹  , Gbadebo A. Ojikutu² , Larry Abdullai¹ ,
Shola Oyedeji¹ , and Jari Porras¹

¹ Department of Software Engineering, LUT University, 53850 Lappeenranta, Finland
mikhail.adisa@lut.fi

² Faculty of Business and Law, Coventry University, Coventry, UK
ojikutug2@uni.coventry.ac.uk

Abstract. Amidst the evolving crises and disruptions threatening firms' competitiveness, businesses are faced with increased dynamism necessitated by technological development, digitalization, and sustainability requirements for survival and growth. This study delves into the intersection of dynamic capabilities (DC), digital transformation (DT), and sustainable resilience among law firms in developing countries. With Nigerian law firms as our case study, this research investigates the strategic integration of dynamic capabilities and digital transformation to foster long-term sustainability of law firms' resilience during a crisis. Through empirical analysis and qualitative exploration, the study unveils obstacles ranging from digital resistance to technical constraints yet uncovers valuable insights from adopting innovative digital strategies that enhance operational resilience and contribute to driving positive economic, environmental, and social impact while ensuring long-term sustainability objectives. The study reaffirms the significance of dynamic capabilities for digital transformation and contributes to the broader discourse on how digital technology enables firms in emerging economies to maneuver disruptions during crises.

Keywords: Dynamic capabilities · Digital transformation · Sustainability · Business resilience · Law firms

1 Introduction

Digital transformation refers to the adoption of innovative digital technologies, including mobile, artificial intelligence, cloud, blockchain, and the Internet of Things (IoT), to significantly enhance business operations, elevate customer experience, and facilitate the creation of novel business models [1]. According to Gobble [2], digitalization typically involves the reconceptualization of entire business processes with the help of digital technology, which culminates in the core integration of digital structures in a new digital business model. Beyond digitalization, the requirements for sustainability commitment and practices have also brought about an increased dynamism to most firms, which further creates novel opportunities for competitive innovation and resiliency [3].

© The Author(s) 2024

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBP 500, pp. 478–492, 2024.

https://doi.org/10.1007/978-3-031-53227-6_34

A digital transformation propelled by digital technologies and dynamic capabilities is typically to gain a competitive advantage and directly create positive business impacts and resiliency [4]. Such transformation can change an entire business model, including, for instance, business communications and intricate internal and external processes, given its unique value-creation process and methods of modification of organizational tasks while fulfilling firms' sustainability goals [4]. For knowledge-intensive business services like those preferred by law firms, for instance, it has been found that digitalization could generally enhance their overall performance [5]. Indeed, research has demonstrated that during the Covid-19 pandemic, the results of firms that were quick to adopt digital transformations were generally positive. For instance, Guo et al. [6] found that digitalization contributed to the improvement of the performance of SMEs during the global pandemic.

Digital sustainable transformation and dynamic capabilities are critical strategic decisions and processes adopted by firms, beginning from the reconceptualization of existing business models and culminating in the remodeling and development of new digital business models to keep businesses afloat and contribute to competitive advantage. The body of literature on the discourse agrees that firms' digitalization of business processes and the integration of dynamic capabilities during a global crisis (for example, Covid-19) largely demonstrated positive impacts on their businesses [6–8]. Our study approaches the research through the theoretical lens of dynamic capabilities, which asserts that a firm's capacity to continuously sense environmental changes, mobilize resources to address them, and transform its operations confers an ability to adapt to emerging crises. Following the framework, we relied on Teece [9, 10] and Yeow et al. [11], who contend that sensing opportunities, seizing them, and flexibly reconfiguring operations through leadership and resource allocation that engage all functions are key to achieving digital transformation during global crises as well as the idea of Zimmer et al., [4] that digital transformation should adopt a digital-sustainable co-transformation perspective focusing on innovations that align with sustainability goals by treating digital transformation and sustainability inseparable components of business strategy and operations for maximum strategic benefit.

Following the existing literature, we identified an essential research gap in that there is no evidence concerning the association of DC, sustainable DT, and law firms' resilience, especially in emerging economies such as Nigeria. Additionally, the study attempts to identify how existing resources, internal processes, and external stakeholders influence sustainable digital transformation among law firms during the global crisis. This study aims to fill the gaps by bringing insight into how Nigerian law firms built on their dynamic capabilities and digital transformation readiness to navigate the global crisis and achieve their sustainability goals successfully. As such, we asked the following research questions to help us investigate the phenomenon:

RQ1: What are the challenges faced by law firms in emerging economies during the Covid-19 pandemic?

RQ2: How did Nigerian law firms utilize dynamic capabilities for sustainable digital transformation during a global crisis?

RQ3: What are the impacts of digital transformation on the sustainable resilience of Nigerian law firms during a global crisis?

To address these research questions, we have collected data in two phases. Phase 1 is an open-ended survey, while Phase 2 involves in-depth interviews. We conducted a qualitative analysis of the data we collected. Our findings highlight the drivers and challenges of adopting digital transformation among Law firms in emerging economies and how the transition impacted their business operations and overall sustainability resilience and goals.

The remaining of this study is organized as follows. Section 2 presents related works on sustainable digital transformation (DT), dynamic capability (DC), Nigerian law firms, and their sustainability goals. It is followed by the description of the empirical data collection and the research process in Sect. 3. Section 4 presents the results and discussions, and Sect. 5 concludes the study.

2 Related Studies

Digital transformation is at the core of redefining a firm's value propositions, leading to a new firm identity, with technology as a central catalyst [12]. It also contributes to firms' sustainability goals [4]. Researchers have emphasized that dynamic capabilities offer unique opportunities for firms to remain competitive over time in an era of environmental dynamism by reconfiguring their resources and capabilities to match and create positive market change [13, 14].

2.1 Sustainable Digital Transformation Amidst Crisis

The advancement in digital technologies has significantly transformed how we live, conduct business activities, and address climate change through digital transformation [15]. At the core of digital transformation initiatives lies a firm's capabilities. Following the emergence of the Covid-19 pandemic, which constitutes a global crisis affecting several businesses and services across the globe, researchers have highlighted how firms responded through digital transformations [6, 16, 17]. Accordingly, firms that were quick to adopt digital transformations during the crisis period significantly improved the quality of their service delivery, improved business operations, and drastically reduced their negative environmental impact [6]. Similarly, business efficiency was enhanced by adopting virtual meetings, virtual offices, and social communications [16, 18] to strengthen brand awareness and engage customers. Another research emphasized the flexibility produced and the development of new, critical technical skills through digitalization processes during the pandemic [19].

Ragazou et al. [16] investigated the evolution of digital transformation in enterprises during the pandemic and discovered that emerging technologies such as blockchain, IoT, artificial intelligence, and machine learning have begun integrating enterprises into their business models. Essentially, organizations were transforming their business models into digital models to accommodate the new circumstances and the overwhelming need for integrating digital technology into their business processes. However, according to Reuschl et al. [17], because of the speed of implementation of digital technology by firms during the pandemic, some organizations were left with limited time to remodel their structures, processes, and cultures in alignment with the new digital environment.

In a similar light, many researchers argue that digital transformation and its efforts are not always successful, including when they are launched during a crisis [20, 21]. In fact, Kochetkov et al. [21] specifically demonstrated that a key challenge associated with the implementation of digital transformation in businesses is that it is not always effective. According to them, this emphasizes the need for firms to conduct prior research into the mode and method of digitalization to assess the possibility or otherwise of the quality and success of their digitalization endeavors [21]. Other challenges may emerge in terms of cost implications and strategic, organizational, cultural, or managerial forms [22]. The changes in organizational structures, strategy, and processes occasioned by technical platforms and big data, given their frequently complex systems and frameworks, can pose serious threats to digitalization efforts, wherein the introduction of members of staff and customers to unfamiliar methods may become hectic and resisted by those who have yet to acclimatize to new technology [19]. These notwithstanding, several researchers have stated that the dynamic capabilities of firms can support their digitalization processes despite extraneous, inhibiting factors such as those presented by the global pandemic crisis [6, 7], including in developing countries [8].

2.2 Dynamic Capabilities of Firms During the Crisis

The dynamic capability theory, rooted in the resource-based theory of a firm, underscores the idea that certain capabilities and resources are difficult to replicate as they constitute unique attributes that serve as the foundation for a firm's competitive advantage. Dynamic capabilities (DC) refer to the comprehensive abilities of firms to develop, integrate, and reconfigure internal and external resources to accelerate adaptation to a rapidly evolving environment to gain competitive advantage and sustainability [9, 10] by creating good opportunities for firms to unleash the potential of their digital DCs.

In the context of crises, DC has three dimensions – sensing capabilities, seizing capabilities, and reconfiguring resources to adapt to the crisis [9, 23]. Sensing capabilities, as used here, underscore the dynamic capability of a firm to recognize threats and/or opportunities from its external business environment [9]. Firms with dynamic capabilities can sense, assess, and understand crises timeously [9, 23]. Although no organization could predict the onset of a global crisis, early assessments could have provided awareness and insights, empowering firms with the data to re-strategize their business processes [6]. Sensing opportunities and threats is fundamental to organizational strategy, especially in a crisis. When firms are aware of potential business threats, they are more likely to identify new opportunities in a given crisis [9].

When firms are equipped with dynamic capabilities, they are more likely to elicit from their external environment information capable of changing their conditions in a crisis [23]. Guo et al. [6] noted an example of the new digital business models launched to solve the challenges associated with contactless delivery during the pandemic. The crisis itself was an opportunity to discover and develop new business models. After successfully seizing capabilities, organizations can recalibrate to judiciously select technologies to re-design their business models [9] and continuously renew organizational routines to ensure alignment. This is referred to as 'reconfiguring resources' in the DC dimensions, and it ensures that firms maintain their survival and competitiveness during crises [9]. The global pandemic outbreak triggered survival instincts among firms, given the high levels

of market uncertainty that propel firms to identify threats and opportunities, understand their positions in the market, and reconfigure their business models accordingly [7]. Overall, dynamic capabilities are critical for the survival of firms in times of crisis and have improved the chances of firms' sustainable resilience during the Covid-19 pandemic [6–8].

2.3 The Nigeria Legal Firms, Digital Technology, and Sustainability Goals

The legal industry in Nigeria boasts over 140,000 lawyers distributed across the federation and actively engaged in the practice of law and who possess expertise cutting across diverse areas of human endeavor [24]. Globally, the legal services industry is a robust interdisciplinary domain, traditionally conservative and often slow to identify innovative technology's capabilities to enhance service delivery [25]. However, it has been discovered that many lawyers now use technology to digitalize and automate monotonous processes, leading to improved productivity and efficiency, eliminating duplication, and enhancing transparency and accountability [25], thereby reducing excessive paperwork and outdated working practices of senior legal professionals. In Nigeria, there has also been a rise in technology adoption in legal practice [26]. Software and digital technologies now aid and support lawyers and judges in executing their daily tasks [26].

Through digital transformation, sustainability is vital for contemporary businesses to gain a competitive advantage, attract customers, and strengthen partnerships incorporating sustainable practices to drive innovation [27]. Additionally, digital technology has proven to be at the forefront of promoting inclusion, resilience, and sustainable development goals in Nigeria by offering a formidable platform that helps to mitigate disruptions that are associated with global crises, such as the Covid-19 pandemic, drive inclusive economic growth, and sustainable development goals [28]. Lawyers use digital apps, electronic mail, and office productivity software daily. Some law firms have subscribed to software that tracks internal and external processes, e-discovery, e-filing processes, smart contracts, alternative dispute resolution, virtual offices, virtual meetings, and virtual court hearings [26, 29]. Digital tools have also helped to significantly lower their carbon footprints, especially from travel and the amount of paper generated yearly.

Furthermore, the Covid-19 pandemic facilitated advancement in the Nigerian legislative system, wherein the adoption of digital technology in legal practice and the courts was put to law as the Court of Appeal Rules were amended to permit the electronic filing of notices of appeal, electronic service via email, and virtual hearings of appeals through audio-visual platforms [30]. In all, software and digital technologies were significantly visible in processes like legal analytics, process automation, scheduling, document management, case management, time management, billing, dispute resolutions, and digital archiving, etc., which improved overall service delivery, eliminated errors, improved turnaround times, customer satisfaction, legal research, reduction in physical commuting, paper wastage, energy consumption, and overall recalibration of resources previously associated with service delivery [26, 29, 30].

3 Methodology

The design selected for this research was the case study design [31]. This design, peculiar to qualitative studies, provides a framework within which a particular case, such as a person, group, event, organization, or industry, is studied (within specific contexts/over specific issues). It generates an in-depth understanding and exploration of real-world complex issues within their natural contexts [31]. We aimed to understand how Nigerian law firms utilize their dynamic capabilities in the sustainable digital transformation of their business processes and the overall impact of such endeavors, especially during a global crisis. We adopted open-ended surveys and semi-structured interviews to collect data from legal professionals to answer the research questions in-depth and within the given case study environment - Nigerian law firms.

3.1 Data Collection Method

We have collected data in two phases. At first, we collected data through an open-ended questionnaire and following the guidelines of Schuler et al. [32], who defined an open-ended survey questionnaire as an efficient method of gathering data from specific groups of respondents. Secondly, we interviewed selected Legal professionals to align and validate findings. Interviews are usually used to collect detailed insights and perspectives concerning social phenomena as they provide an excellent platform for collecting rich, contextual data to formulate theories in inductive reasoning. They are indispensable for many qualitative studies, including case studies [31, 33]. Using the purposive sampling techniques, we identify and select appropriate participants for the study. Purposive sampling is a kind of qualitative sampling adopted to specifically select participants who fall within the requisite category for research [33].

The survey (N = 14) and the interviews (N = 18) were conducted specifically for legal professionals (senior associates, senior managers, practice managers, and managing partners) and chief technology officers (CTOs) in law firms with head offices in Lagos, Nigeria. Each participant has up to 10 years and above experience in the industry, except for the senior associate, who has less than 10 years of experience. The size of the firm they represented was between 50 to over 100 employees. The rationale behind the emphasis on this sample population was informed by their involvement in driving their firms' sustainability goals and business process transformation. A participant invitation/consent letter was initially sent to 45 targeted participants via email and other digital means, such as WhatsApp, but only 22 honored our request.

Additionally, we asked the 22 engaged participants for referrals to deepen our data collection, which resulted in 10 extra willing participants. The data collection took place for two months (October and November 2022). The interview sessions were conducted via Zoom and Microsoft Teams and were recorded with the participant's consent. The survey was designed using Google Forms. Altogether, 32 participants distributed across 12 law firms were involved in the survey and the interview. Table 1 gives a summary of their demographic distributions.

Table 1. Excerpt of analytical memo table displaying our coding process and strategies.

No	Job role	Interview participants	Years of experience	Job role	Survey participants
1	Managing partner	4	Over 10 years	Managing partner	3
2	Practice manager	3	Over 10 years	Practice manager	2
3	Senior manager	1	Over 10 years	Senior manager	3
4	CTO	8	Over 10 years	CTO	2
5	Senior associates	2	5 – 10	Senior associates	4
	Total	18			14

3.2 Data Analysis Method

We adopted thematic analysis as the preferred data analysis technique for this study due to its appropriateness in identifying, evaluating, and reporting themes, categorizations, patterns, areas of convergence, and divergence within the data [34]. After recording the interviews using the audio recorders, the researcher transcribed them using Otter.Ai, a voice-and-video-to-text transcription and analysis software. Next, we selected an appropriate coding strategy to enable us to identify relevant information called empirical indicators and code them [35]. Coding was done manually using the Microsoft Visio application to encourage a deeper involvement with the data and accurate interpretation and construction. Three researchers were involved in coding and categorizing the data from the surveys and interviews process, as shown in Table 2. The results from the two phases were merged, including data relating to the same firm or question after a repeated and careful analysis to arrive at the final thematic schemes reflecting the research questions.

Table 2. Excerpt of analytical memo table displaying our coding process and strategies.

Anchor code	Interview and Survey	Empirical indicator	Extraction of key indicator	Coding and categorization	Thematic scheme
RQ 1 to 3 distributions	Responses from the survey and interview were tabulated and analyzed	Merging of similar responses from the survey interview	Keywords were extracted for mapping	Keywords were categorized and mapped to RQs	Final mapping to 1st to 3rd-order themes

4 Result

This section covers the data analysis findings concerning the research questions.

4.1 Findings and Discussion

The study reveals three core themes. Firstly, the challenges of law firms during the pandemic and their barriers in DT transition efforts. Secondly, how DC factors and DT readiness helped them to overcome the barriers and challenges, and Thirdly, the impact of their DC and DT efforts on the sustainability goals and sustainable resilience of their business. The resulting themes from the triangulation of findings from open-ended surveys and interviews are presented in thematic coding (see Fig. 1). We discussed the findings in relation to the research questions for the emerging descriptive, second-order, third-order, and core themes. The resulting impacts were represented as positive (+) and negative (-) signs, respectively.

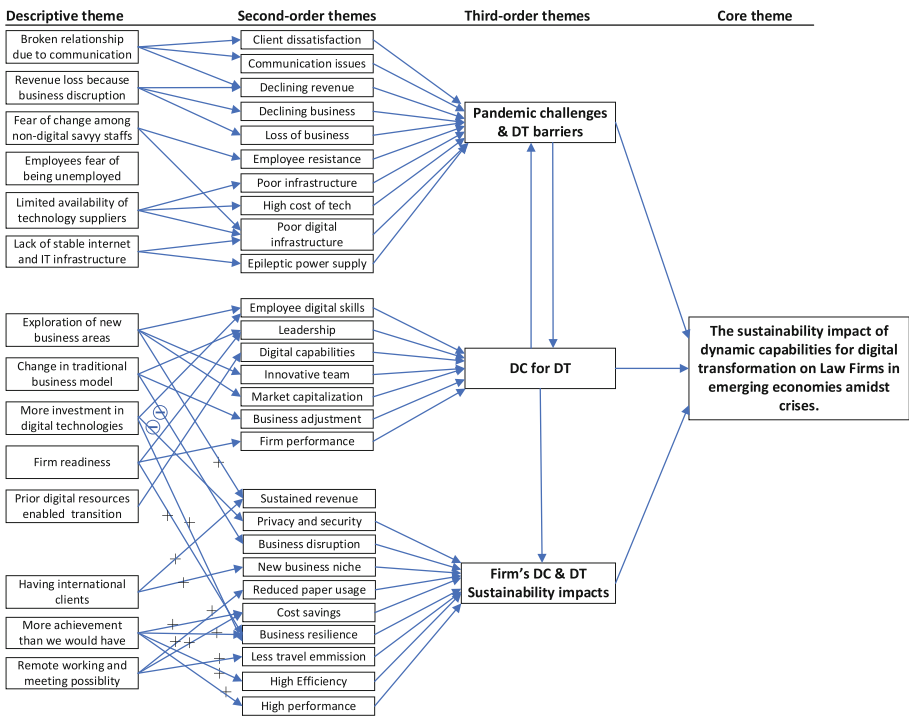


Fig. 1. Thematic coding of digital transformation in Nigeria law firms

Utilization of Dynamic Capabilities for Digital Transformation During a Global Crisis. The insights from our research reveal similar results to [36] that sensing, seizing, and reconfiguring elements of dynamic capability are key to achieving digital transformation during global crises like the Covid-19 pandemic in developing countries. Our findings indicate that Nigerian law firms effectively utilized digital dynamic capabilities during the crisis by sensing and seizing the opportunities in the digital space and reconfiguring their digital resources to continuously adapt internal structures and processes to remain competitive as the digital landscape evolves. Thus, the firms were able to

leverage digital technologies to enhance sensing, seizing, and transformation to improve their sustainable resilience and overall service delivery and management.

Sensing and Seizing Opportunities in the Digital Space. Our study revealed that while the crises disrupt the law firms' businesses, it create simultaneous opportunities for those with existing resources who have demonstrated prior commitment to sustainable practice and technology. Most of our respondents envisaged the opportunities of digital business transformation and being seen as a promoter of sustainability practices and have been gradually investing and improving their digital infrastructure and lowering their environmental impacts. Others claimed that external factors and the DT trends within the judicial and other sectors in developed countries influenced their DC. Quoting a CTO on how his firm sense and seize the opportunities in the digital space, *"We were using digital means before then; we just had to explore it further and see what we could achieve by proceeding with the transformation because we were clear about the impact, of that transformation. And again, it wasn't when we were mindful that there could be glitches along the way. But I guess with much determination, knowing the outcome we desired, we were positive through the proof."*

Furthermore, other participants indicated that they conducted research and consulted with technology experts concerning whether digital assets would help seize the new market opportunity and overcome the threats posed by the pandemic. For example, a Managing partner stated, *"The first point of call was our IT personnel...what is accessible to our clients? We conducted an internal staff survey, which revealed that we could sustain. And by relying on technology, we can transform our processes and still reduce carbon footprints simultaneously."*

Reconfiguring Digital Resources. Many participants confirmed they had developed new internal policies and processes, supported activities, and organized training to facilitate digitalization efforts and strengthen their sustainability goals. Most firms introduce new digital policies, adopt hybrid internal operations, restructure their strategies, and invest in communicating them. In addition, a 'pro-environmental culture awareness campaign' forms a significant part of their strategy. Enforcement of duplex printing, reduced paper waste, digital archiving, and email signatures for all outgoing emails were introduced to remind staff and clients to consider the environment before printing emails. A Practice manager said, *"Our entire business model changed from normal brick and mortar. Yeah, it has changed. Now, we're investing much more in digitalized resources and services. As expected, there were kickbacks and dissenting views... And there were trainings ... We looked for key stakeholders who we believe can drive a vision to other team members. And that's how we particularly spread the goodwill."* Furthermore, another CTO said, *"Well, yes, we had to align our processes and modify our technology policy... things like analytics and cybersecurity became very key, we had to take training on basic cybersecurity... two-factors authentication, screen lock, how to keep your document, how to keep your computers, you know, because we all work in virtually and later hybrid and we are concerned with our information, as well as client information... many things became digital... we now have a lot of virtual meetings, even our training."* Insights from our findings indicate that DC is a critical success factor for sustainable DT in law firms as firms with strong DC were able to reconfigure their resources judiciously to re-design their business processes and continually align their organizational routines.

A changing business model and developing new routines, processes, policies, trainings, etc., drives Nigerian law firms to adapt to the changing business climate while reducing their environmental impact.

Challenges and Barriers. In delivering the digital transformation processes, most of the law firms highlight human resistance to change, late technology adoption, staff skillsets, broken relationships, finances, epileptic power supply, weak infrastructure, network and bandwidth, organizational strategies, and overall costs of digital transformation as major challenges faced during the digital transformation initiatives. Importantly, while the younger staff members demonstrated early commitment, the older staff expressed many reservations at the beginning of the process. For example, another Practice manager stated, *“There was an increase in the budget allocated for digital information technology. It is more than double the previous budget.”* Similarly, a Senior associate confirmed, *“We do experience poor network connection while connecting to the office server due to weak internet network where we lived, wherein some staff had to resort to using multiple data sources to access office resources.”*

Sustainability Impacts of Digital Transformation on Nigerian Law Firms During Crisis. The findings reveal that the adoption of digital dynamic capabilities had the following effects on Nigerian law firms: Recombining multiple digital assets to support new and existing business processes was achieved through adopting and integrating digital assets, accessibility, leadership, effective stakeholder management, and long-term planning. An improved performance was achieved through enhanced efficiency, improved firm output, and business resilience, but with noticeable differences among the firms’ reconfiguration of internal and external resources. The above confirmed that law firms’ investment in digital technology and sustainability practices significantly impacts their transformation and sustainable resiliency.

Enhanced Adoption and Integration of Digital Assets. We deduced that firms with huge financials and investment in digital technology footprints could seamlessly enhance and transform into digitally enabled law firms than those with less financial capability. This leads to the maximization of their resources to improve efficiency and productivity and streamline communication processes with their clients. For example, a Senior manager revealed, *“During the pandemic, a lot didn’t change for us, besides moving from a physical location to working remotely and later hybrid, it was seamless for our teams. Digitalization enhanced our international and local operations. Because we were prepared, we have invested in legal software and technologies to support our operations, we have always been pro-environmental in service delivery and dealt with international clients.”*

Accessibility and Effective Stakeholder Management. Our findings revealed that accessibility and effective stakeholder management became more feasible with digital technology, resulting in enhanced business processes and communication with both clients and partners. The accessibility was facilitated by various software solutions deployed across the firms, which allow their client to access digital records, update documents, and track the progress of their legal assets or cases without leaving their home or offices. An excerpt from a Senior associate: *Accessibility was crucial during*

the crisis. Everybody was on their laptop and cell phones, working remotely, accessing centralized resources, and assuring our clients of our robust service delivery.”

Long-Term Planning and Enhanced Efficiency. Most of our respondents agreed that the tendency for digitalization to support future initiatives and service delivery is enormous. The technology-facilitated achievements during the crisis have all become a normal operation procedure for most law firms after the crisis. This was evident in improved performance, service delivery, cost reduction, valuable analytics, efficient tracking, and pro-environmental consciousness facilitated by the available digital technologies. Another Managing partner responded - *“As far back as 2016 and 2017, we were already moving digital. We could see that, oh, foreign firms have been using legal software and having different meetings remotely with partners in Nigeria (via Skype and conference calls), and this is how they’re doing it. So, it’s more like we could spy into the future and then draw us into the future.”* This finding confirms that digital transformation has the propensity to enhance firm performance and can aid in planning business processes.

Improved Firm Output and Sustainable Resilience. Another discovery in this study was the improved firm output due to digitalization, emphasized by many respondents. The law firms navigated through the crisis successfully and achieved a new level that seemed unreachable before the crisis. They were able to reduce their environmental impact through the deployed technology infrastructure, as most activities that were previously done manually and on paper have now been digitized. Digital archiving and other pro-environmental activities become the norm. Overall, the technology kept the firms afloat throughout the crisis and beyond. These firms are today competing favorably with their foreign counterpart in driving Nigeria’s legal practices. An excerpt from one of the Senior managers - *“The firm maintained its usual excellence, performance, and service delivery to clients. Our client grew, our data expanded sporadically, our technology budget increased. But we have the results: we saved time to commute, shortened response time, gave more access to customers, and increased productivity because we now have digital solutions and tools.”*

Our findings indicate that digital resources are essential for business survival and competitive advantage, which aligns with results from [37]. Furthermore, digitalization can improve firm output, productivity, and performance, especially for firms in knowledge-intensive business services, like law firms. Despite these positive impacts, there is evidence of a few unintended negative impacts on the firms. First, DT disrupted the business model of the law firms. They had to change their long-standing business traditions and learn to use new technologies and software. It caused much resistance, especially among staff who were not tech-savvy (often among senior employees).

Consequently, this led to a digital divide where those with prior knowledge of technology quickly adapted while others were left behind. In addition, digitalization could expose law firms to cybersecurity attacks and ransomware, thus requiring additional infrastructure procurement. Before digitalization, client files were kept in hard copies under locks and keys that were not easily accessible by unauthorized people. However, digitalization could expose people’s privacy, especially if the firm does not have a strong cyber security team and software.

4.2 Threat to Validity

Our research is subject to threats to validity, including internal, external, and conclusion validity. The threats to the study's validity and mitigation [38] are discussed for completeness. Internal validity relates to a causal relationship. The participants were recruited based on their experience, knowledge, and positions from different law firms without being coarse. Their responses and experiences differ from each other. However, the credibility of their responses was enhanced by triangulation comparing the survey and interview responses to form thematic codes validated by all the authors while maintaining an independent standpoint, keeping an open mind, and acting in good faith throughout the study. External validity relates to generalizing our findings across multiple industries and settings. All our participants are from Law firms in Nigeria. Thus, the findings of this study are not generalizable.

Given that the findings of qualitative studies are not generalizable due to their highly contextualized nature. Thus, the findings of this study are not generalizable. However, the research methods may be adopted to study the same or a similar phenomenon in other case settings and contexts [33]. Conclusion validity relates to the degree to which conclusions drawn from the relationships in data are reasonable. The participants were grouped into two sets to compare and validate responses from multiple participants with different experience levels and involvement in the digitalization processes. These produced a database for making the right judgments concerning the transferability of the findings.

4.3 Research Limitations

The findings of this study are only relevant to law firms in emerging economies like Nigeria. Another research limitation may have been the inability to get a wider sample size as initially planned. The fear of releasing firms' strategies prevented others from honoring our request. We also observed that some participants could have been biased in their responses. However, we believed their responses as professional practitioners.

A future study may aim for a broader sample size within ethical limits. A further limitation may have been the virtual conduction of the interviews. Given the nature of qualitative studies, it is ideal to conduct research in natural environments and observe body movements and gesticulations to support the interpretation of data, etc. [33]. However, the researcher's engagement with the respondents and the manual coding enhanced the validity and reliability of the study.

5 Conclusion

In conclusion, Nigerian law firms, like many other businesses across the globe, were not immune to crises. The Covid-19 pandemic impacted the business processes of Nigerian law firms adversely through revenue decline, occasioned by restrictions on business activities, shortage of business opportunities and force majeure, breakdown in physical interaction, and overwhelming uncertainty and fear. In their quest to digitally transform their business operations, they faced barriers and challenges such as employee resistance

to change, lack of digital infrastructure, and unreliable power supply. However, building on their dynamic capabilities, they were able to reconfigure their business operations and discover new business opportunities for survival, competitiveness, and the overall sustainability of their business.

Adopting dynamic capabilities resulted in investments in digital transformation and strengthened by visionary leadership, resulting in sustainable resilience of their business and positive economic and environmental impacts. Findings from the study indicate that Nigerian Law firms' efficiency, performance, revenue, and business resilience improved tremendously. Furthermore, they were able to save costs on energy, transportation, and printing, as well as improve the working conditions of employees. However, DT also resulted in unintended consequences such as privacy, security, and business disruption. Although business disruption has eventually become the new normal, privacy and security issues are something Nigerian Law firms will continue to invest in, just as many companies around the world would have to deal with in the digital economy era.

References

1. Warner, K.S.R., Wäger, M.: Building dynamic capabilities for digital transformation: an ongoing process of strategic renewal. *Long Range Plann.* **52**, 326–349 (2019). <https://doi.org/10.1016/J.LRP.2018.12.001>
2. Gobble, M.A.M.: Digitalization, digitization, and innovation. *Res. Technol. Manage.* **61**, 56–59 (2018). <https://doi.org/10.1080/08956308.2018.1471280>
3. Linde, L., Sjödin, D., Parida, V., Wincent, J.: Dynamic capabilities for ecosystem orchestration a capability-based framework for smart city innovation initiatives. *Technol Forecast Soc Change.* **166**, 120614 (2021). <https://doi.org/10.1016/j.techfore.2021.120614>
4. Zimmer, M.P., Järveläinen, J.: Digital–sustainable co-transformation: introducing the triple bottom line of sustainability to digital transformation research. In: Kreps, D., Davison, R., Komukai, T., Ishii, K. (eds.) *Human Choice and Digital by Default: Autonomy vs Digital Determination*. HCC 2022. IFIP Advances in Information and Communication Technology, vol. 656, pp. 100–111. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15688-5_10
5. Ribeiro-Navarrete, S., Botella-Carrubi, D., Palacios-Marqués, D., Orero-Blat, M.: The effect of digitalization on business performance: an applied study of KIBS. *J. Bus. Res.* **126**, 319–326 (2021). <https://doi.org/10.1016/J.JBUSRES.2020.12.065>
6. Guo, H., Yang, Z., Huang, R., Guo, A.: The digitalization and public crisis responses of small and medium enterprises: implications from a COVID-19 survey. *Front. Bus. Res. China* **14**, 1–25 (2020). <https://doi.org/10.1186/S11782-020-00087-1/FIGURES/1>
7. Martins, A.: Dynamic capabilities and SME performance in the COVID-19 era: the moderating effect of digitalization. *Asia-Pacific J. Bus. Admin.* **15**, 188–202 (2023). <https://doi.org/10.1108/APJBA-08-2021-0370>
8. Owoseni, A., Hatsu, S., Tolani, A.: How do digital technologies influence the dynamic capabilities of micro and small businesses in a pandemic and low-income country context? *Electron. J. Inf. Syst. Dev. Countries.* **88**, e12202 (2022)
9. Teece, D.J.: Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strateg. Manag. J.* **28**, 1319–1350 (2007)
10. Teece, D.J.: Business models and dynamic capabilities. *Long Range Plann.* **51**, 40–49 (2018). <https://doi.org/10.1016/J.LRP.2017.06.007>

11. Yeow, A., Soh, C., Hansen, R.: Aligning with new digital strategy: a dynamic capabilities approach. *J. Strateg. Inf. Syst.* **27**, 43–58 (2018). <https://doi.org/10.1016/J.JSIS.2017.09.001>
12. Ulfesnes, R., Moe, N.B., Hanssen, G.K., Buan, T.A.: interrelation of digitalization and digital transformation in a maritime company. In: Carroll, N., Nguyen-Duc, A., Wang, X., Stray, V. (eds.) *Software Business, ICSOB 2022, LNBIP*, vol. 463, pp. 51–66. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-20706-8_4
13. Teece, D.J., Pisano, G., Shuen, A.: Dynamic capabilities and strategic management. *Strateg. Manag. J.* **18**, 509–533 (1997)
14. Eisenhardt, K.M., Martin, J.A.: Dynamic capabilities: what are they? *Strategic Manag. J. Strat. Mgmt. J.* **21**, 1105–1121 (2000)
15. Dwivedi, et al.: Climate change and COP26: are digital technologies and information management part of the problem or the solution? An editorial reflection and call to action. *Int. J. Inf. Manage.* **63**, 102456 (2022). <https://doi.org/10.1016/J.IJINFOMGT.2021.102456>
16. Ragazou, K., Passas, I., Sklavos, G.: Investigating the strategic role of digital transformation path of SMEs in the era of covid-19: a bibliometric analysis using R. *Sustainability* **14**, 11295 (2022). <https://doi.org/10.3390/SU141811295>
17. Reuschl, A.J., Deist, M.K., Maalaoui, A.: Digital transformation during a pandemic: stretching the organizational elasticity. *J. Bus. Res.* **144**, 1320–1332 (2022). <https://doi.org/10.1016/J.JBUSRES.2022.01.088>
18. Jacobides, M.G., Cennamo, C., Gawer, A.: Towards a theory of ecosystems. *Strateg. Manag. J.* **39**, 2255–2276 (2018). <https://doi.org/10.1002/SMJ.2904>
19. Aljanazrah, A., Yerousis, G., Hamed, G., Khlaif, Z.N.: Digital transformation in times of crisis: challenges, attitudes, opportunities and lessons learned from students’ and faculty members’ perspectives. *Front Educ. (Lausanne)*. **7**, 1047035 (2022). <https://doi.org/10.3389/FEDUC.2022.1047035/BIBTEX>
20. Budagov, A.S., Sukhova, N.A.: Problems of effective business digital transformation management. *Eur. Proc. Soc. Behav. Sci.* 428–434 (2020)
21. Kochetkov, E.P., Zabavina, A.A., Gafarov, M.G.: Digital transformation of companies as a tool of crisis management: Ð N empirical research of the impact on efficiency. *Strat. Decis. Risk Manag.* **12**, 68–81 (2021)
22. Nahrkhalaji, S.S., Shafiee, S., Shafiee, M., Hvam, L.: Challenges of digital transformation: the case of the non-profit sector. In: *IEEE International Conference on Industrial Engineering and Engineering Management*. 2019-December, pp. 1245–1249 (2019)
23. Ballesteros, L., Useem, M., Wry, T.: Masters of disasters? An empirical analysis of how societies benefit from corporate disaster aid. *AMJ* **60**, 1682–1708 (2017). <https://doi.org/10.5465/AMJ.2015.0765>
24. The NBA: who we are. <https://nigerianbar.org.ng/>. Accessed 25 Aug 2023
25. Michalakopoulou, K., Nikitas, A., Njoya, E.T., Johnes, J.: Innovation in the legal service industry: Examining the roles of human and social capital, and knowledge and technology transfer. *Int. J. Entrep. Innov.* (2022)
26. Joshua Olujobi, O.: COVID-19 pandemic and evolving trends in legal practice in Nigeria: challenges and prospects. *Run Law J.* **4**, 111–126 (2020)
27. Martínez-Peláez, R., et al.: Role of digital transformation for achieving sustainability: mediated role of stakeholders, key capabilities, and technology. *Sustainability* **15**, 11221 (2023). <https://doi.org/10.3390/SU151411221>
28. Benjamin, O.O., Foye, O.V.: Inclusion, organizational resilience, and sustainable development in Nigeria: the role of digital innovations. *Environ. Sci. Proc.* **15**, 27 (2022)
29. Arinze-Umobi, C.N., Oknonkwo, I.T.: Alternative dispute resolution practice in Nigeria and the effect of COVID-19 pandemic. *Int. J. Law Clin. Legal Educ.* **2**, 82 (2021)

30. Mathew, U.: Making legal technology a viable sector In Nigeria (2021). <https://www.mondaq.com/nigeria/fin-tech/1142658/making-legal-technology-a-viable-sector-in-nigeria>. Accessed 25 Aug 2023
31. Yin, R.K.: Case study research design and applications. *Can. J. Program Eval.* 319 (2018)
32. Schluter, J., Seaton, P., Chaboyer, W.: Critical incident technique: a user's guide for nurse researchers. *J. Adv. Nurs.* **61**, 107–114 (2008). <https://doi.org/10.1111/J.1365-2648.2007.04490.X>
33. Bell, E., Harley, B., Bryman, A.B., Bell, E., Alan, B., and B.H.: *Business Research Methods*. Oxford University Press, Oxford (2022)
34. Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qual. Res. Psychol.* **3**, 77–101 (2006). <https://doi.org/10.1191/1478088706QP063OA>
35. Adu, P.: *A step-by-step guide to qualitative data coding*. Routledge, New York (2019)
36. Putritamara, J.A., Hartono, B., Toiba, H., Utami, H.N., Rahman, M.S., Masyithoh, D.: Do dynamic capabilities and digital transformation improve business resilience during the covid-19 pandemic? Insights from beekeeping MSMEs in Indonesia. *Sustainability* **15**, 1760 (2023). <https://doi.org/10.3390/SU15031760>
37. Bai, C., Quayson, M., Sarkis, J.: COVID-19 pandemic digitization lessons for sustainable development of micro-and small-enterprises. *Sustain Prod. Consum.* **27**, 1989–2001 (2021)
38. MacCoun, R., Shadish, W.R., Cook, T.D., Campbell Boston, D.T.: *Experimental and quasi-experimental designs for generalized causal inference*. Cengage Learn. (2002). <https://doi.org/10.1016/J.SPC.2021.04.035>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.





Research Streams of Barriers to Digital Transformation: Mapping Current State and Future Directions

Henning Brink¹ , Fynn-Hendrik Paul¹  , and Sven Packmohr² 

¹ BOW, Osnabrück University, Osnabrück, Germany
{henning.brink, fynn-hendrik.paul}@uos.de

² DVMT, Malmö University, Malmö, Sweden
sven.packmohr@mau.se

Abstract. Digital Transformation (DT) strives to alter an entity by substantially changing its characteristics facilitated by integrating digital technologies. Albeit numerous barriers hinder the realization of its potential. Barriers are subject to scientific research. Generally, scientific works result in research streams. The existing literature already examines the DT streams. Although these works make an essential contribution, they cannot sufficiently explore the field of barriers. Keeping track of the concepts and themes in a growing research field is challenging. Therefore, the aims of this mapping study are (1) to show which domain-specific research streams are explicitly dealing with the DT barriers, (2) to highlight which topics research currently addresses, and (3) which topics should be investigated in the future. Combining elements of a bibliometric analysis with a systematic literature review, we mapped nine different streams based on 203 publications. The results indicate that much research focuses on industrial companies or sectors but needs an overarching perspective. Also, many studies are only concerned with identifying the barriers, while systematic approaches to overcoming them still need to be developed.

Keywords: Digital Transformation · Barriers · Research Streams · Mapping Study · Literature Review

1 Introduction

Digital technologies profoundly impact society, the economy, and daily life [1]. Digital transformation (DT), characterized by significant changes through information, computing, communication, and connectivity technologies, promises micro, meso, and macro benefits. It influences how individuals work and spend their free time [2]. At the meso level, businesses can experience improved efficiency, productivity, and revenue [3], leading to higher living standards at the macro level [3]. Organizations often face barriers when attempting to fully leverage the transformative potential of digital technologies [4]. DT encompasses integrating digital technologies, leading to socio-technical changes

within organizations [1, 5]. Barriers, derived from innovation management and organizational change research, hinder or prevent DT activities [6, 7]. Barriers are factors “that can hinder or stop the successful implementation of DT” [8]. Research has predominantly focused on success factors [9]. However, since barriers are more than the mere opposite of success factors, the results cannot simply be transferred [10]. Understanding these barriers is crucial for effective implementation and requires identification, analysis, and appropriate countermeasures. Previous studies on barriers have primarily focused on digitalization rather than the broader scope of DT [11, 12]. Thus, they cannot grasp the scope and scale of DT, which requires additional in-depth research [12, 13]. Luckily, researchers are increasingly examining barriers in the context of DT. However, as this research field is increasingly growing, keeping track of the different concepts and themes is getting challenging. The growing field of barriers in DT research necessitates comprehensive exploration to capture diverse concepts and themes [4]. This study aims to identify the research streams and topics related to DT barriers. Mapping studies have arisen to help fulfill this aim. These studies aim to review “a relatively broad topic by identifying, analyzing, and structuring the goals, methods, and contents of conducted primary studies” [14]. In comparison, while a “conventional systematic literature review makes an attempt to aggregate the primary studies in terms of the research outcomes [...], a mapping study usually aims [...] to classify the relevant literature” [15]. Mapping studies identify broader topics such as research streams, their central subject areas, and untreated areas. [14] Mapping studies are, therefore, particularly valuable as they provide a foundation for future research [15]. Thus, our research questions are as follows: What are the research streams in the field of barriers to digital transformation? Which topics are addressed within the research streams? What research needs have been outlined within the research streams?

The study is structured as follows: First, we introduce the topic and give a brief theoretical background. After, we present the methodology of our data collection. The results comprise different clusters found in the literature and give an aggregate view of current studies and their views on future research. We close with a concluding discussion.

2 Theoretical Background

With the rapid advancements in digital technologies and their increasing impact on various aspects of society and business, the term “digital transformation” emerged. There are multiple definitions for the term available in the literature. Based on various definitions of DT, Vial constructed a conceptual definition of DT as a significant alteration of an entity’s characteristics through the integration of information, computing, communication, and connectivity technologies, utilizing new digital technologies [1]. Gong and Ribere unified DT as “a fundamental change process enabled by digital technologies that aims to bring radical improvement and innovation to an entity [e.g., an organization, a business network, an industry, or society] to create value for its stakeholders by strategically leveraging its key resources and capabilities” [16]. These definitions clearly distinguish DT from other related terms. While digitization primarily focuses on converting analog information into digital form, and digitalization pertains to the adoption of digital technologies in specific processes, DT has a comprehensive socio-technical

impact on the entire organization [1, 11]. The scope of DT even goes beyond terms like IT-enabled organizational transformation (ITOT). In contrast to ITOT, DTs redefine the value proposition of organizations and create new organizational identities, while ITOT revolves around supporting the existing value proposition and reinforcing the organizations' identity by leveraging digital technologies [12]. Consequently, regarding DT, all departments within an organization are affected and must navigate changes such as the adoption and implementation of new digital technologies, processes, structures, and potential financial barriers [4].

In recent years, many researchers in information systems have therefore studied concepts, impacts, and aspects of DT from a variety of perspectives [1]. One field of research examines the barriers to DT. However, research on barriers did not start in the context of the DT. The research field builds on areas such as innovation management [5] and organizational change [13]. Transferred from the field of innovation research, a barrier is defined as "an issue that either prevents or hampers" [14] DT activities in an organization. Due to DT, socio-technical structures previously mediated by non-digital relationships and artifacts are transformed to be mediated by digital relationships and artifacts [15]. The tensions that arise from this integration of physical and digital layers are named barriers to DT [16]. Examining barriers is essential as they differ from success factors [10]. Even though success factors are the earlier research concept, they evolved into barriers as their understanding is vital for effective implementation [13].

3 Method

Our mapping study aims to provide an overview of research on DT barriers. We combine bibliometric analysis elements with a systematic literature review to achieve this aim. Our qualitative and quantitative approaches can be divided into 3 phases.

Phase 1 (Development of the search strategy and database selection): We discussed possible search terms to identify literature related to our research topic. We decided on using the search string "(Digital Transformation) AND Barrier", as other terms like "digitalization" do not capture the essence of the subject under investigation. The Scopus database was chosen because it contains a wide range of scientific literature and allows exporting search hits, which is necessary for our bibliometric analysis.

Phase 2 (Carrying out the literature search and selecting literature): Applying the search string, we got 374 hits in November 2022. Only English-language, peer-reviewed scientific literature from journals or conference proceedings was considered. We explicitly excluded articles whose research focus was not related to DT barriers. Following the recommendations of vom Brocke [17], we examined the hits' titles, abstracts, and keywords to check for relevance. We identified 171 entries without relation to our subject matter, leaving us with 203 relevant publications.

Phase 3 (Analysis of the Literature): The last phase is separated into a quantitative and qualitative literature analysis. We performed the quantitative analysis with techniques of bibliometric analysis. Beginning with a performance analysis, we analyzed the most important metrics of the research, such as the number of publications per year and citations. These metrics assess the productivity and impact of a research field [18]. Afterward, we conducted science mapping to investigate the relationship between the

research articles. We analyzed the author and index keywords using VOSViewer and the co-occurrence [17] feature to derive research streams. The co-occurrence or “co-word analysis assumes that words that frequently appear together have a thematic relationship with one another” [18]. Thus, we obtained different thematic clusters consisting of various keywords using VOSViewer. Compared to a purely manual subjective sorting of research articles, applying a co-word analysis can determine given word correlations exploratively, quantitatively, and objectively [19]. However, as word usage can vary between specific and general [18], we discussed the thematic clusters and their keywords among the authors. We manually refined the topics in these discussions by aggregating and reassigning keywords. Combining both approaches allowed us to minimize their disadvantages. The results of this phase are nine distinguishable thematic clusters representing research streams. Afterward, we continued the analysis using qualitative content analysis [20]. We read every publication and assigned each publication to one stream. Conducting an open coding approach within a group of individual researchers, we marked relevant phrases describing the research objectives and research outlook. By applying the analytical induction [21], we merged similarities to set up topics. For each stream, we could then understand which topics are currently being investigated and which should be investigated in the future.

4 Results

The dataset includes a total of 203 publications spanning 11 active years. These publications involve contributions from 637 authors, demonstrating a diverse and collaborative research environment. Among the publications, 19 were solely authored, while 183 resulted from collaborative efforts. The average productivity per active year of publication is calculated to be 22.44, indicating a consistent output of research within the field. The collaboration index, calculated to be 0.016, suggests a relatively low level of collaboration among authors within the field. However, the collaboration coefficient of 0.68 indicates a moderate degree of collaboration, as most publications result from collaborative efforts. The number of publications steadily increased from one publication in 2015 to two publications in 2016, and further increased in 2017 (4), 2018 (13), 2019 (32), and 2020 (37). In 2021, 61 publications were recorded, followed by 51 publications in 2022. These variations in publication numbers suggest fluctuations in research activity and focus within the field during the examined period. The total number of citations received by the publications amounts to 2757, with an average of 14 citations per publication and 306 per year. Out of the total publications, 137 were cited, representing 67.82% of the overall publications. Results indicate that the publications in this field have acquired significant attention and impact within the scholarly community. Following our research approach, we identified nine different research streams, as shown in Table 1. In the following, the streams are presented. To make our findings more transparent, we exemplarily reference selected studies we identified.

The stream of **Industry 4.0** addresses a range of research aims to identify, measure, and overcome barriers associated with Industry 4.0 and Internet of Things (IoT) implementation. Publications consider specific industrial environments like manufacturing, farming, food, and electronics. With eight publications, supply chains and their

Table 1. Research streams of barriers to digital transformation.

Research Stream (Number of publications) Keywords	Main research focus	Main research needs
<p>Industry 4.0 (44) Additive manufacturing, industrial research, industrial revolutions, industry 4.0, internet of things, smart manufacturing, human-computer interaction, supply chains, supply chain management, interoperability, productivity, quality control, value creation</p>	<ul style="list-style-type: none"> • Barriers associated with implementation of IoT • Industrial environments such as manufacturing, farming, food, and electronics • Impacts on supply chain management • Readiness and practices of SMEs • Conditions of the COVID-19 pandemic 	<ul style="list-style-type: none"> • Empirical research and real case scenarios • Focusing on more sectors • Conducting longitudinal studies of SMEs • Refining maturity models
<p>Technology Adoption (33) Innovation, digital innovation, technological innovation, technology, digital technologies, emerging technologies, modern technologies, information and communication technology, information technology, information systems, intelligent systems, adoption, technology adoption, integration, artificial intelligence, blockchain, machine learning, 3D printers, social media, enterprise social media, big data, data analytics, security of data</p>	<ul style="list-style-type: none"> • Barriers associated with adoption of digital technologies such as blockchain and artificial intelligence • Data-related topics like cybersecurity, big data, and data governance • Strategies for overcoming adoption barriers 	<ul style="list-style-type: none"> • Empirical research • Extending the geographical, sectoral, and organizational scope • Applying quantitative research approaches • Investigating interrelationships among barriers

(continued)

Table 1. (continued)

Research Stream (Number of publications) Keywords	Main research focus	Main research needs
<p>Service Industry (22) Electronic commerce, international trade, banking, commerce, e-commerce</p>	<ul style="list-style-type: none"> • Barriers hindering logistics service providers, retail, banking, legal services, and cultural heritage management • Strategies for overcoming barriers • Impacts on customer relationships, revenue management, and supply chain risk management 	<ul style="list-style-type: none"> • Exploring the role of digital platforms, emerging digital technologies, and digital ecosystems • Conducting comparative studies across industries and sectors • Larger sample sizes • Applying multi-case and -method approaches • Developing strategies for overcoming barriers, especially SMEs
<p>Education (22) Curricula, digital devices, digital inclusion, e-learning, engineering education, higher education, learning systems, personnel training, students, teaching, digital skills, digital literacy</p>	<ul style="list-style-type: none"> • Perspectives of different stakeholder groups on barriers to DT in education institutions • Mainly individual institutions, specific countries, or subject areas such as management • Conditions of the COVID-19 pandemic • Strategies for overcoming adoption barriers 	<ul style="list-style-type: none"> • Larger sample sizes • Conducting comparative studies across different types of education institutes • Exploring contextual and sociodemographic factors influencing the perception of barriers • Collaboration among researchers, educators, and industry professionals to advance the field

(continued)

Table 1. (continued)

Research Stream (Number of publications) Keywords	Main research focus	Main research needs
<p>Public Sector (21) Digital government, e-government, information use, public administration, public organizations, smart city, European Commission, non-profit organization</p>	<ul style="list-style-type: none"> • Barriers to the shift from governments to digital or smart governments • DT of public administrations • Role of governments in causing regulatory barriers and in overcoming barriers • Non-profit organizations 	<ul style="list-style-type: none"> • Conducting studies across different countries • Larger sample sizes • Applying quantitative research approaches • Investigating correlation between DT process and barriers • Conducting longitudinal studies • Understanding differences between organizational- and individual-level barriers • Developing strategies for overcoming barriers
<p>Management (20) Business model innovation, business modeling, digital platforms, knowledge transfer, managers, enabling technologies, industrial management, risk assessment, change management, organizational change, decision making, strategy, digital strategy, information management, innovation management, investments</p>	<ul style="list-style-type: none"> • Managerial barriers associated with implementation of digital technologies • Recommendations for action for managing the DT process • Agile project management • Digital Entrepreneurship 	<ul style="list-style-type: none"> • Conducting comparative studies across different industries and organizational processes • Exploring methods and actions to overcome barriers • Role of competences of a workforce for the DT process and overcoming barriers

(continued)

Table 1. (continued)

Research Stream (Number of publications) Keywords	Main research focus	Main research needs
Construction (13) Construction industry, architectural design, building information modeling	<ul style="list-style-type: none"> • Potential and adoption of digital technologies such as production robots, 3D-printing, and BIM software • Different countries such as Germany, South Africa, and North Macedonia • Role of different stakeholder groups for DT process 	<ul style="list-style-type: none"> • Exploring factors influencing the adoption of digital technologies • Exploring strategies and best practices for overcoming barriers • Empirical research • Applying quantitative research approaches
Healthcare (10) Health care delivery, e-health, digital health, telehealth, telemedicine, human, medical informatics, data privacy, covid-19	<ul style="list-style-type: none"> • Barriers associated with adoption of monitoring technology or health apps • Perspectives of different stakeholder groups on digital technologies 	<ul style="list-style-type: none"> • Conducting comparative studies across different countries, cultures, healthcare areas, stakeholder groups, and technologies • Applying mixed-method and quantitative research approaches
Residual (18) Digital economy, economics, economic and social effects	<ul style="list-style-type: none"> • Barriers to DT in the energy sector, rural areas, accounting, and small businesses 	<ul style="list-style-type: none"> • Larger sample sizes and in-depth analysis • Applying mixed-method and quantitative research approaches

management are one of the key areas of research. Researchers analyze how DT affects procurement processes and their integration into supply chain operations. Research also identifies major barriers hindering the adoption of digital supply chain practices and analyzes their interrelationships. Additionally, the stream concentrates on the readiness and practices of small and medium-sized enterprises (SMEs) in adopting Industry 4.0 either holistically [22–24] or with a regional focus [25–28]. Surveys are conducted to assess the readiness of IoT or Industry 4.0 adoption. Furthermore, the stream includes publications analyzing barriers to DT during the COVID-19 pandemic [29]. Case studies and projects are examined to understand the current status and future prospects of Industry 4.0 implementation. Frameworks and methodologies for DT beyond traditional approaches are proposed to guide companies on their DT journeys.

Regarding further research, the majority of publications do not suggest concrete further research approaches. However, publications state that empirical research and real case scenarios are needed to understand the barriers to the implementation of Industry 4.0, e.g., in sustainability-focused supply chains [30] or manufacturing processes [31]. Research should focus on more sectors beyond just manufacturing [32]. Bertello et al. [33] emphasize the need to monitor SMEs over a longer period of time. In this regard, Ghobakhloo et al. [22] formulate research questions on how SMEs should prioritize approaches to adopting Industry 4.0 technologies and which competence sets SMEs should develop in this context. Furthermore, publications show the need for research to refine maturity models to assess the companies' status quo and the effectiveness of DT projects. Herceg et al. [32] propose maturity models considering DT holistically by including a broader range of dimensions, such as culture and leadership. With a more holistic perspective in the context of manufacturing, but confirming the previous proposals for future research, some scholars develop research agendas for any dimension of their specifically developed barrier model [8]. These agendas comprise examples of research questions for the barrier dimensions of missing skills, technical barriers, individual barriers, organizational and cultural barriers, and environmental barriers.

The **Technology Adoption** stream encompasses studies exploring the potential and barriers associated with adopting and implementing new technologies in different industries and organizational contexts like SMEs. The study's primary objective is to uncover and analyze the factors that hinder or facilitate the integration of these technologies and propose strategies for successful DT. To do so, they are based on literature but also on case studies and surveys. A prominent area of investigation within this stream focuses on the adoption and utilization of blockchain, e.g., in manufacturing [34, 35] and supply chains [36]. These studies aim to identify the potential benefits of blockchain adoption while also analyzing the barriers incumbent companies face in leveraging this technology effectively. Another key aspect of the stream involves studying the impact and operationalizing of artificial intelligence in general [37] or in specific use cases like robotic process automation [38] or container management for smart manufacturing [39].

Data-related topics like cybersecurity, big data, and data governance also form important areas of investigation. Studies present conceptual frameworks and propose solutions to enhance organizations' cybersecurity approaches and data governance systems. In addition, studies aim to understand the requirements and use of big data. In terms of future research directions, the majority of publications do not give a precise research

outlook. However, researchers recommend empirical studies that extend the geographical, sectoral, and organizational scope. Furthermore, Flechsig et al. [38] propose to apply quantitative research approaches to validate and complement previous findings. Moreover, Vafadarnikjoo et al. [34] emphasize investigating the interrelationships among identified barriers and other factors.

One major topic in the *Service Industry* stream is the identification of barriers hindering DT in various service industries [40], such as logistics service providers, cultural heritage management, retail, banking, and legal services. Thus, exploring DT's drivers [41], as opposed to barriers, influencing digitalization efforts in different sectors, including luxury hotels, sub-Saharan Africa's financial inclusion, B2B companies, and leading banks, seems a valid research strategy. Other scholars provide insights into successful strategies, leading practices, and organizational elements contributing to effective DT in diverse contexts, such as logistics providers [13], retail operations, and museums' communication strategies. The investigation of the impact of DT on customer relationships, revenue management, and supply chain risk management. Especially in service industries, innovative digital approaches to navigating external contingencies like the COVID-19 pandemic seem crucial [42]. These approaches might be used in e-Commerce adoption [43] as well as the implementation of banking services.

Based on the studies' suggestions, future research should explore the role of digital platforms, emerging technologies (e.g., blockchain, AI, IoT), and digital ecosystems in industries like logistics [13], hotels, and banking. Investigating their impact on performance, competitiveness, revenue management, and customer behavior will provide actionable insights. Additionally, developing measurement scales for evaluating the intangible aspects [44] of brand awareness and customer engagement is crucial. Conducting comparative studies across industries and sectors will identify common challenges and opportunities in DT [13]. Examining the influence of different contexts, such as geography, culture, and organizational characteristics, will provide valuable strategies for diverse settings. Larger sample sizes and multi-case, multi-method approaches will enhance generalizability and validity [45]. Research should focus on understanding and addressing barriers to successful DT. Developing adaptable implementation strategies, especially for small organizations [46], will be valuable. Examining the impact of regulations on digital technologies, mobile banking, social media [42], and omnichannel implementation will guide policymakers and organizations.

Studies in the stream of *Education* include the perspectives of different stakeholder groups, such as students, teachers, and academic and administrative staff, on barriers to DT in education institutions. Schools, as well as public and private universities, are examined. The data are usually based on an individual university or a specific country. Cross-national studies, such as from Eri et al. [47], are rare. In addition, some studies focus on specific subject areas, such as management [48]. The majority of studies present a list or model of identified barriers. The studies are partly influenced by the COVID-19 pandemic or explicitly address the impact of the pandemic [48]. Literature reviews summarize these barriers [49]. Some studies also present recommendations for overcoming barriers [50]. Aditya et al. further aimed at developing a framework for identifying, assessing, and prioritizing barriers, as the "existing literature has reported

a barrier list that could affect the implementation of DT in higher education, yet the research question of how to identify barriers remained unanswered” [51].

Regarding research outlooks, many publications recommend an expansion of the database [49]. Studies should aim to validate the results with a more diverse stakeholder group to include different perspectives [48], and explore contextual and sociodemographic factors influencing the perception of barriers [48, 52]. A stronger collaboration among researchers, educators, and industry professionals is emphasized to advance the field [53]. Research is needed to compare barriers in different higher education types [49], to understand how they relate to each other and how they could be overcome [54].

Most papers in the stream *Public sector* examined barriers to the shift from governments to digital or smart governments [55] or the DT of public administrations [56]. While many studies have addressed barriers to DT within these settings, there have also been studies that have examined the role of governments in causing regulatory barriers [57] or their role in overcoming barriers, e.g., for small service businesses [58]. A few studies deal with the barriers to DT in non-profit organizations [59], also in comparison to for-profit organizations [60]. Ablyazov and Ungvári [61] identified barriers in the smart city context. Compared to the “Healthcare” stream, relatively few papers address specific technologies, such as cloud computing adoption for government services [62].

Future research in this field could include several countries [56] or a large number of organizations in their database “in order to be able to generalize the results” [63]. Quantitative Studies to validate “in various and broader contexts” [64] are advised as with other streams. Studies like these could examine the correlation between the DT process and the barriers [56] or examine the changes over time by performing longitudinal studies [65]. Also, research on a better understanding of the differences between organizational-level and individual-level barriers is recommended [66]. Again, more research on overcoming barriers is called a research outlook [66].

The stream *Management* focuses on understanding and addressing the opportunities and barriers that organizations and managers encounter when implementing digital technologies. In summary, the publications aim to provide recommendations for action for managing the DT process. The stream emphasizes the importance of managing structural changes and removing organizational barriers influencing the transformation process. The publications address special topics: agile project management [67, 68] and digital entrepreneurship [69]. Except for one publication dealing with the banking sector [70], the stream does not contain sectoral references.

In terms of further research, this stream emphasizes providing insights for managers and organizations navigating the challenges of DT in the future. Studies recommend investigating different industries and organizational processes to improve the understanding of how different methods and actions can be used to overcome barriers to DT. Additionally, Ciampi et al. [67] propose to explore the impact of digital competences on the relationship between DT and organizational agility. Biclesanu et al. [69] suggest cross-country comparisons to broaden the observations and generalize the findings.

Studies in the stream *Construction* examine the construction industry in different countries such as Germany, South Africa, and North Macedonia. Some focus on the benefits of DT, such as case studies of production robots, 3D printing, and BIM software [71]. Scholars advocate for digital partnering in South Africa’s construction industry

based on a survey of construction professionals. The study explores how interactions among architects, clients, contractors, and consultants shape industry characteristics and options for DT [72]. Further studies evaluate BIM adoption, emphasizing barriers in technology and management. Also, opportunities for integrating BIM into education are discussed [73]. Scholars identify barriers to DT in architecture using organizational learning theory. Barriers to DT, such as missing adoption of data-centric approaches or AI-enhanced sensor networks in construction. Finally, scholars research decision-making for end-of-life facilities to promote sustainable practices [74].

Further research should encompass understanding the factors influencing the adoption and successful implementation of digital technologies in construction, such as their drivers, barriers, and enablers. This includes exploring strategies for overcoming resistance to change and identifying best practices for effective adoption [75]. Research should involve developing comprehensive frameworks and methodologies for assessing aspects such as productivity, cost efficiency, sustainability, safety, and quality. At best with quantitative analysis, case studies, and comparative evaluations. Further research on integrating emerging digital technologies, such as artificial intelligence, robotics, augmented reality, and blockchain, is needed to foster innovation in the construction industry [71]. Also, organizational factors such as leadership styles, cultural aspects, change management strategies, collaboration models, and communication approaches need further attention for successful digital partnering and collaboration [73].

Research in the *Healthcare* stream strongly focuses on technologies, such as monitoring technology [76] or health apps [77]. Poncette et al. [78] examine the barriers to integrating new technologies that are limited to intensive care units. Based on the technology focus of the studies, a large majority of the studies survey the users of the technologies, particularly doctors, nurses, and other clinical staff [79]. Natsiavas et al. [80] examined how citizens feel about sharing their health data with healthcare professionals or eHealth providers. As in other streams, most articles focus on identifying barriers.

In this stream, many studies recommend broadening the data base in future research, e.g., by including more countries to identify cultural differences [79] or more stakeholders, such as patients [79] or the management of healthcare organizations [81]. Further research should also consider environmental characteristics such as the physical environment, the nature of the department, and organizational policies [81]. Several studies also recommend greater validation of results through mixed-method studies [79] or additional quantitative results [81]. The studies in this stream mostly focus on individual areas or technologies, lacking an overall holistic and socio-technical view of an organization. In the “Healthcare” stream, research is needed that applies a comprehensive view of DT as a combination and integration of different digital technologies to improve an organization by triggering significant changes [1].

Residuals cover papers that did not fit into the other streams or covered singular aspects, such as DT in the energy sector, rural areas, or the perceptions and challenges of DT in accounting [82]. Another singular aspect is public sector adaptation to enhance improved service delivery and organizational resilience [83]. Other studies explore barriers to IoT in water management, hinders in small businesses regarding

blockchain, or factors stimulating/inhibiting Smart Grid development [84]. Examining cross-cultural barriers in DT highlights technology's potential in diverse business environments. More generally, some papers examine barriers and enablers of DT. One proposes a socio-technical model categorizing barriers [85].

Future work is suggested by further analyzing living labs and rural stakeholders' context to identify driver barriers and impact patterns. Co-designing a system and developing requirements for citizen involvement is necessary [83]. In accounting, research should focus on the impact of digitalization and the role of public entities [82]. Investigating resistance to change, culture, and price as barriers is crucial. For digitalization in the energy sector, research should explore managerial barriers and evaluate opportunities, risks, and competencies [84]. Especially for generic models, larger samples and in-depth analysis are needed. Research involves collecting quantitative data, using mixed-methods approaches, and adapting models as digitalization evolves [85].

5 Concluding Discussion

This mapping study has provided a comprehensive analysis of the research streams. The identified streams offer a holistic understanding of the multifaceted nature of this research field and provide a foundation for future studies in this field. Our findings indicate a strong thematic focus on private-sector companies. The underlying reasons for this can be multifaceted. Due to the strong economic importance or their impact on society, this sector might be in the spotlight. Industry frequently serves as a leading example, e.g., achieving efficiency gains, adapting to evolving work dynamics, and exploring diverse avenues for value creation [86]. The unequal distribution could be related to DT's advancement, data availability, research funding, or research interests. This has different implications for research in the field of DT barriers. Looking at the different streams in comparison helps identify gaps in less advanced streams. In addition, the degree to which findings can be transferred should be examined. Collaboration among researchers from different disciplines and industries could provide new insights.

Although the streams differ regarding their themes, certain commonalities can be observed regarding the research approaches in the studies. It is striking that most of the studies adopt a qualitative approach. Quantitative and mixed-method approaches, by contrast, are much rarer. The high proportion of qualitative studies could be related to the relatively young age of the research field and the short publication period of most studies starting from 2019. For a research field with little pre-existing knowledge, qualitative research is better suited to gain new insights compared to quantitative approaches [87]. In the light of model development phases [88], most publications are in the phase of designing the models, respectively identifying the barriers. Research must now address "how this can be measured" [89]. Then, scholars need to test and evaluate the models to assess their reliability, validity, and generalizability [88, 89]. Measurement instruments and procedural models can also help practitioners to identify and prioritize barriers in specific real-world scenarios [51]. Research also needs to develop recommendations for overcoming barriers effectively. Barriers could become facilitators if they are mastered [9]. A wider use of quantitative approaches would also allow the examination of the relationship between barriers and other constructs, such as the DT process or financial

metrics. In addition, which factors have an influence on the perception of barriers could be investigated. Noticeable, however, is a lack of a clear research outlook in many scientific articles. A clear research outlook is essential for guiding future research efforts and identifying emerging trends and challenges within the field. Researchers should strive to provide a concise but explicit research outlook in their articles, highlighting the areas for further investigation.

The implications of our study are manifold. Our study provides an overview of the research efforts in the field and guides scientists in their future research. The study also offers implications for practitioners who want to embrace DT. It allows them to get a quick and systematic overview of the current body of knowledge and evidence in the field of barriers to DT. The streams related to industries especially allow practitioners to better identify barriers and help accelerate the DT process, e.g., how to develop and implement strategies or what corporate culture and competencies are advantageous. Further, for academics, the more general streams can serve as a broader perspective in driving research programs forward. Also, our work identifies underrepresented streams and topics of future interest, serving as a foundation for formulating funding programs.

However, it is essential to acknowledge the limitations of this study. As the research field continually evolves, stream changes will likely occur over time. By combining a bibliometric with a systematic literature review, we attempted to counterbalance the disadvantages of each method to derive the streams objectively. However, it is still possible that other scientists will reach a different outcome through different inferences or methods. Further, the restriction to the Scopus database and the inclusion and exclusion criteria used to select relevant literature may have influenced the results.

Our mapping study has already revealed several research needs, which are presented in the result section. Regarding research on the streams in barriers to DT, we can further note that future research should focus on exploring specific research streams in greater detail to provide more nuanced insights. Periodic reviews should be conducted to determine how the research field is changing. Further research could also include perspectives from practitioners and industry to derive a more comprehensive research agenda.

References

1. Vial, G.: Understanding digital transformation: a review and a research agenda. *J. Strateg. Inf. Syst.* **28**, 118–144 (2019)
2. Selimović, J., Pilav-Velić, A., Krndžija, L.: Digital workplace transformation in the financial service sector: Investigating the relationship between employees' expectations and intentions. *Technol. Soc.* **66**, 101640 (2021)
3. Corejova, T., Chinoracky, R.: Assessing the potential for digital transformation. *Sustainability.* **13**, 11040 (2021). <https://doi.org/10.3390/su131911040>
4. Hess, T., Matt, C., Benlian, A., Wiesböck, F.: Options for formulating a digital transformation strategy. *MIS Q. Exec.* **15**, 123–139 (2016)
5. Verhoef, P.C., et al.: Digital transformation: A multidisciplinary reflection and research agenda. *J. Bus. Res.* **122**, 889–901 (2021). <https://doi.org/10.1016/j.jbusres.2019.09.022>
6. Sandberg, B., Aarikka-Stenroos, L.: What makes it so difficult? A systematic review on barriers to radical innovation. *Ind. Mark. Manage.* **43**, 1293–1305 (2014)

7. Yoo, Y., Lyytinen, K.J., Boland, R.J., Berente, N.: The next wave of digital innovation: opportunities and challenges: a report on the research workshop digital challenges in innovation research. *SSRN J.* (2010). <https://doi.org/10.2139/ssrn.1622170>
8. Vogelsang, K., Liere-Netheler, K., Packmohr, S., Hoppe, U.: Barriers to digital transformation in manufacturing: development of a research agenda. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*, pp. 4937–4946 (2019)
9. Hadjimanolis, A.: The Barriers Approach to Innovation. In: Shavinina, L.V. (ed.) *The International Handbook on Innovation*, pp. 559–573. Pergamon, Oxford (2003)
10. Williams, J., Ramaprasad, A.: A taxonomy of critical success factors. *Eur. J. Inf. Syst.* **5**, 250–260 (1996). <https://doi.org/10.1057/ejis.1996.30>
11. Cardinali, S., Pagano, A., Carloni, E., Giovannetti, M., Governatori, L.: Digitalization processes in small professional service firms: drivers, barriers and emerging organisational tensions. *J. Serv. Theory Pract.* **33**, 237–256 (2022)
12. Pellathy, D.A., In, J., Mollenkopf, D.A., Stank, T.P.: Middle-range theorizing on logistics customer service. *IJPDLM.* **48**, 2–18 (2018)
13. Cichosz, M., Wallenburg, C.M., Knemeyer, A.M.: Digital transformation at logistics service providers: barriers, success factors and leading practices. *IJLM.* **31**, 209–238 (2020)
14. Wendler, R.: The maturity of maturity model research: a systematic mapping study. *Inf. Softw. Technol.* **54**, 1317–1339 (2012)
15. Kitchenham, B.A., Budgen, D., Pearl Brereton, O.: Using mapping studies as the basis for further research – a participant-observer case study. *Inf. Softw. Technol.* **53**, 638–651 (2011). <https://doi.org/10.1016/j.infsof.2010.12.011>
16. Gong, C., Ribiere, V.: Developing a unified definition of digital transformation. *Technovation* **102**, 102217 (2021). <https://doi.org/10.1016/j.technovation.2020.102217>
17. vom Brocke, J., et al.: Reconstructing the giant: on the importance of rigour in documenting the literature search process. In: *ECIS 2009 Proceedings*. (2009)
18. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M.: How to conduct a bibliometric analysis: an overview and guidelines. *J. Bus. Res.* **133**, 285–296 (2021). <https://doi.org/10.1016/j.jbusres.2021.04.070>
19. Rha, J.S., Lee, H.-H.: Research trends in digital transformation in the service sector: a review based on network text analysis. *Serv. Bus.* **16**, 77–98 (2022)
20. Mayring, P.: Qualitative content analysis. *Forum qualitative sozialforschung/forum: qualitative social research. Qual. Methods Various Disciplines I: Psychol.* (2000). <https://doi.org/10.17169/FQS-1.2.1089>
21. Glaser, B.G., Strauss, A.L.: *Discovery of Grounded theory: Strategies for qualitative research*. Routledge, Milton Park (2017)
22. Ghobakhloo, M., Iranmanesh, M., Vilkas, M., Grybauskas, A., Amran, A.: Drivers and barriers of Industry 4.0 technology adoption among manufacturing SMEs: a systematic review and transformation roadmap. *J. Manuf. Technol. Manag.* **33**, 1029–1058 (2022). <https://doi.org/10.1108/JMTM-12-2021-0505>
23. Stentoft, J., Jensen, K.W., Philipsen, K., Haug, A.: Drivers and Barriers for Industry 4.0 Readiness and Practice: A SME Perspective with Empirical Evidence. Presented at the Hawaii International Conference on System Sciences (2019)
24. Bettiol, M., Capestro, M., Maria, E.D., Micelli, S.: SMEs @ Industry 4.0: a comparison between top and average performers. *Sinergie Italian J. Mgmt.* **39**, 27–48 (2021)
25. Jankowska, B., Mińska-Struzik, E., Bartosik-Purgat, M., Götz, M., Olejnik, I.: Industry 4.0 technologies adoption: barriers and their impact on polish companies' innovation performance. *Eur. Plan. Stud.* **31**, 1029–1049 (2023)
26. Bucci, G., Bentivoglio, D., Finco, A., Belletti, M.: Exploring the impact of innovation adoption in agriculture: how and where precision agriculture technologies can be suitable for the Italian farm system? *IOP Conf. Ser.: Earth Environ. Sci.* **275**, 012004 (2019)

27. Mezentseva, E.: Digital transformation of Russian industry: the specifics of large, medium and small enterprises. In: Kumar, V., Leng, J., Akberdina, V., and Kuzmin, E. (eds.) *Digital Transformation in Industry*. pp. 307–319. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-94617-3_22
28. Peillon, S., Dubruc, N.: Barriers to digital servitization in French manufacturing SMEs. *Procedia CIRP*. **83**, 146–150 (2019). <https://doi.org/10.1016/j.procir.2019.04.008>
29. Kutnjak, A.: Covid-19 accelerates digital transformation in industries: challenges, issues, barriers and problems in transformation. *IEEE Access*. **9**, 79373–79388 (2021)
30. Machado, E., Scavarda, L.F., Caiado, R.G.G., Thomé, A.M.T.: Barriers and Enablers for the Integration of Industry 4.0 and Sustainability in supply chains of MSMEs. *Sustainability*. **13**, 11664 (2021). <https://doi.org/10.3390/su132111664>
31. Rahnama, H., Johansen, K., Larsson, L., Rönnbäck, A.Ö.: Collaboration in value constellations for sustainable production: the perspective of small technology solution providers. *Sustainability*. **14**, 4794 (2022). <https://doi.org/10.3390/su14084794>
32. Vuksanović Herceg, I., Kuč, V., Mijušković, V.M., Herceg, T.: Challenges and Driving Forces for Industry 4.0 Implementation. *Sustainability*. **12**, 4208 (2020)
33. Bertello, A., De Bernardi, P., Ricciardi, F.: Open innovation for digital transformation in low- and medium-tech SMEs: analysis of pre-competitive collaborative projects. In: Ceci, F., Prencipe, A., Spagnoletti, P. (eds.) *Exploring Innovation in a Digital World*. LNISO, vol. 51, pp 182–197. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87842-9_14
34. Vafadarnikjoo, A., Badri Ahmadi, H., Liou, J.J.H., Botelho, T., Chalvatzis, K.: Analyzing blockchain adoption barriers in manufacturing supply chains by the neutrosophic analytic hierarchy process. *Ann Oper Res*. (2021). <https://doi.org/10.1007/s10479-021-04048-6>
35. Lohmer, J., Lasch, R.: Blockchain in operations management and manufacturing: Potential and barriers. *Comput. Ind. Eng.* **149**, 106789 (2020)
36. Hackius, N., Petersen, M.: Translating high hopes into tangible benefits: how incumbents in supply chain and logistics approach blockchain. *IEEE Access*. **8**, 34993–35003 (2020)
37. Hajishirzi, R., Costa, C.J.: Artificial Intelligence as the core technology for the digital transformation process. In: 2021 16th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1–6 (2021)
38. Flechsig, C., Anslinger, F., Lasch, R.: Robotic Process automation in purchasing and supply management: a multiple case study on potentials, barriers, and implementation. *J. Purch. Supply Manag.* **28**, 100718 (2022)
39. Kim, B.S., Lee, S.H., Lee, Y.R., Park, Y.H., Jeong, J.: Design and implementation of cloud docker application architecture based on machine learning in container management for smart manufacturing. *Appl. Sci.* **12**, 6737 (2022)
40. Manny, L., Duygan, M., Fischer, M., Rieckermann, J.: Barriers to the digital transformation of infrastructure sectors. *Policy. Sci.* **54**, 943–983 (2021)
41. Bollweg, L., Lackes, R., Siepermann, M., Weber, P.: Drivers and barriers of the digitalization of local owner operated retail outlets. *J. Small Bus. Entrep.* **32**, 173–201 (2020). <https://doi.org/10.1080/08276331.2019.1616256>
42. Feitosa Jorge, L., Mosconi, E., Santa-Eulalia, L.A.: Enterprise social media platforms for coping with an accelerated digital transformation. *JSIT*. **24**, 221–245 (2022)
43. Zhou, Y., Muller, L.M.: Technical, managerial, and organizational changes in adapting an e-commerce model in civil engineering design and consulting services. *Leadersh. Manage. Eng.* **3**, 202–212 (2003)
44. Hazée, S., Zwienerberg, T.J., Van Vaerenbergh, Y., Faseur, T., Vandenberghe, A., Keutgens, O.: Why customers and peer service providers do not participate in collaborative consumption. *JOSM*. **31**, 397–419 (2020). <https://doi.org/10.1108/JOSM-11-2018-0357>
45. Diener, F., Špaček, M.: Digital transformation in banking: a managerial perspective on barriers to change. *Sustainability*. **13**, 2032 (2021). <https://doi.org/10.3390/su13042032>

46. Feitosa Jorge, L., Mosconi, E., Santa-Eulalia, L.A.: Overcoming the barriers of an enterprise social media initiative. In: AMCIS 2020 Proceedings (2022)
47. Eri, R., et al.: Digital resilience in higher education in response to COVID-19 pandemic: Student Perceptions from Asia and Australia. *JUTLP*. **18**, 108–134 (2021)
48. Packmohr, S., Brink, H.: Impact of the pandemic on the barriers to the digital transformation in higher education - comparing pre- and intra-covid-19 perceptions of management students. In: Buchmann, R.A., Polini, A., Johansson, B., and Karagiannis, D. (eds.) *Perspectives in Business Informatics Research*, pp. 3–18. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87205-2_1
49. Aditya, B.R., Ferdiana, R., Kusumawardani, S.S.: The department of electrical and information engineering, Gadjah Mada University, Indonesia, categories for barriers to digital transformation in higher education: an analysis based on literature. *IJJET*. **11**, 658–664 (2021). <https://doi.org/10.18178/ijjet.2021.11.12.1578>
50. Carnicero, I., González-Gaya, C., Rosales, V.F.: The transformation process of the university into a data driven organisation and advantages it brings: qualitative case study. *Sustainability*. **13**, 12611 (2021). <https://doi.org/10.3390/su132212611>
51. Aditya, B.R., Ferdiana, R., Kusumawardani, S.S.: A barrier diagnostic framework in process of digital transformation in higher education institutions. *J. Appl. Res. Higher Educ.* **14**, 749–761 (2021). <https://doi.org/10.1108/JARHE-12-2020-0454>
52. Livingston, E., Houston, E., Carradine, J., Fallon, B., Akmeemana, C., Nizam, M., McNab, A.: Global student perspectives on digital inclusion in education during COVID-19. *Glob. Stud. Child.* **13**(4), 341–357
53. Alstrup, S., Rootzén, H.: Possibilities and barriers for e-learning in primary school in Denmark. In: 15th European Conference on e-Learning. Presented at the (2016)
54. Aditya, B.R., Ferdiana, R., Kusumawardani, S.S.: The study of the barriers to digital transformation in higher education: a preliminary investigation in Indonesia. In: 2020 6th International Conference on Science and Technology (ICST), pp. 1–6 (2020)
55. Wilson, C., Mergel, I.: Overcoming barriers to digital government: mapping the strategies of digital champions. *Gov. Inf. Q.* **39**, 101681 (2022)
56. Tangi, L., Janssen, M., Benedetti, M., Noci, G.: Barriers and drivers of digital transformation in public organizations: results from a survey in the Netherlands. In: Viale Pereira, G., Janssen, M., Lee, H., Lindgren, I., Rodríguez Bolívar, M.P., Scholl, H.J., and Zuiderwijk, A. (eds.) *Electronic Government*, pp. 42–56. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-57599-1_4
57. Akatkin, Y., Yasinovskaya, E.: Data-centricity as the key enabler of digital government: is Russia ready for digital transformation of public sector. In: Chugunov, A., Misnikov, Y., Roshchin, E., and Trutnev, D. (eds.) *Electronic Governance and Open Society: Challenges in Eurasia*, pp. 439–454. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-13283-5_33
58. Chen, C.-L., Lin, Y.-C., Chen, W.-H., Chao, C.-F., Pandia, H.: Role of government to enhance digital transformation in small service business. *Sustainability*. **13**, 1028 (2021)
59. O’Grady, J., Roberts, P.: The digital transformation of Irish non-profit organisations (2019)
60. Vogelsang, K., Packmohr, S., Brink, H.: Challenges of the digital transformation – comparing nonprofit and industry organizations. In: Ahlemann, F., Schütte, R., and Stieglitz, S. (eds.) *Innovation Through Information Systems*, pp. 297–312. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86790-4_21
61. Ablyazov, T., Ungvári, L.: Digital platforms of territory management. In: Jahn, C., Ungvári, L., and Ilin, I. (eds.) *Algorithms and Solutions Based on Computer Technology*, pp. 313–325. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93872-7_26

62. Alenizi, A.S., Al-karawi, K.A.: Cloud computing adoption-based digital open government services: challenges and barriers. In: Yang, X.-S., Sherratt, S., Dey, N., and Joshi, A. (eds.) *Proceedings of Sixth International Congress on Information and Communication Technology*, pp. 149–160. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-1781-2_15
63. Rusu, L., Balasuriya, P.B.L., Bah, O.: Cultural barriers in digital transformation in a public organization: a case study of a Sri-Lankan organization. In: Themistocleous, M., Papadaki, M., and Kamal, M.M. (eds.) *Information Systems*, pp. 640–656. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63396-7_43
64. Wang, Y.-F.: Existence and intentionality of digital transformation in public organizations: a phenomenological perspective. In: *DG.O 2022: The 23rd Annual International Conference on Digital Government Research*, pp. 400–409. ACM, Virtual Event Republic of Korea (2022). <https://doi.org/10.1145/3543434.3543444>
65. Abdullahi, A., Mustafa, A.S.: Barriers to the effective implementation of e-parliament in Africa: a case study on the 8th Nigeria national assembly (2015–2019). *EG*, **17**, 438 (2021)
66. Rösler, J., Söll, T., Hancock, L., Friedli, T.: Value co-creation between public service organizations and the private sector: an organizational capabilities perspective. *Admin. Sci.* **11**, 55 (2021). <https://doi.org/10.3390/admsci11020055>
67. Ciampi, F., Faraoni, M., Ballerini, J., Meli, F.: The co-evolutionary relationship between digitalization and organizational agility: ongoing debates, theoretical developments and future research perspectives. *Technol. Forecast. Soc. Chang.* **176**, 121383 (2022). <https://doi.org/10.1016/j.techfore.2021.121383>
68. Fuchs, C., Hess, T.: *Becoming agile in the digital transformation: the process of a large-scale agile transformation* (2018)
69. Biclesanu, I., Anagnoste, S., Branga, O., Savastano, M.: Digital entrepreneurship: public perception of barriers, drivers, and future. *Admin. Sci.* **11**, 125 (2021)
70. Sund, K.J., Bogers, M.L.A.M., Sahramaa, M.: Managing business model exploration in incumbent firms: a case study of innovation labs in European banks. *J. Bus. Res.* **128**, 11–19 (2021). <https://doi.org/10.1016/j.jbusres.2021.01.059>
71. Berlak, J., Hafner, S., Kuppelwieser, V.G.: Digitalization’s impacts on productivity: a model-based approach and evaluation in Germany’s building construction industry. *Prod. Plann. Control.* **32**, 335–345 (2021)
72. Linderoth, H.C.J., Jacobsson, M., Elbanna, A.: Barriers for digital transformation: the role of industry. In: *ACIS 2018 - 29th Australasian Conference on IS*, pp. 1–11 (2018)
73. Pereiro-Barceló, J., Meléndez, C.: Introducing BIM into education: opportunities and challenges. In: *IV International Conference on Civil Engineering Education* (2018)
74. Nik-Bakht, M., et al.: Value stream mapping of project lifecycle data for circular construction. In: *Proceedings of the International Symposium on Automation and Robotics in Construction*, pp. 1033–1042 (2021)
75. Gardner, N.: Digital transformation and organizational learning: situated perspectives on becoming digital in architectural design practice. *Front. Built Environ.* **8**, 905455 (2022)
76. Dugstad, J., Eide, T., Nilsen, E.R., Eide, H.: Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. *BMC Health Serv. Res.* **19**, 366 (2019)
77. Leigh, S., Ashall-Payne, L., Andrews, T.: Barriers and facilitators to the adoption of mobile health among health care professionals from the United Kingdom: discrete choice experiment. *JMIR Mhealth Uhealth* **8**, e17704 (2020). <https://doi.org/10.2196/17704>
78. Poncette, A.-S., Meske, C., Mosch, L., Balzer, F.: How to overcome barriers for the implementation of new information technologies in intensive care medicine. In: Yamamoto, S.,

- Mori, H. (eds.) *Human Interface and the Management of Information. Information in Intelligent Systems*, vol. 11570, pp. 534–546. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22649-7_43
79. Iyanna, S., Kaur, P., Ractham, P., Talwar, S., Najmul Islam, A.K.M.: Digital transformation of healthcare sector. What is impeding adoption and continued usage of technology-driven innovations by end-users? *J. Bus. Res.* **153**, 150–161 (2022)
 80. Natsiavas, P., Kakalou, C., Votis, K., Tzovaras, D., Koutkias, V.: Citizen perspectives on cross-border ehealth data exchange: a European survey. *Stud Health Technol Inform.* **264**, 719–723 (2019). <https://doi.org/10.3233/SHTI190317>
 81. Leeuw, J.A.D., Woltjer, H., Kool, R.B.: Identification of factors influencing the adoption of health information technology by nurses who are digitally lagging: in-depth interview study. *J. Med. Internet Res.* **22**, e15630 (2020)
 82. Gonçalves, M.J.A., Da Silva, A.C.F., Ferreira, C.G.: The future of accounting: how will digital transformation impact the sector? *Informatics.* **9**, 19 (2022)
 83. Fleron, B., Pries-Heje, J., Baskerville, R.: Digital organizational resilience: a history of denmark as a most digitalized country. In: Presented at the Hawaii International Conference on System Sciences (2021). <https://doi.org/10.24251/HICSS.2021.294>
 84. Kuceba, R., Blok, F.: Digital transformation of energy management in micro smart grid networks. In: Garci-Perez, A., Simkin, L. (eds.) *Proceedings of the 22nd European Conference on Knowledge Mgmt. Academic Conferences International Ltd., Reading* (2021)
 85. Brink, H., Packmohr, S., Paul, F.-H.: Extending a socio-technical model of the barriers to digital transformation through data triangulation. In: *2022 8th International Conference on Information Management (ICIM)*, pp. 68–74. IEEE, Cambridge, United Kingdom (2022). <https://doi.org/10.1109/ICIM56520.2022.00020>
 86. Brynjolfsson, E., McAfee, A.: *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company, New York, NY (2014)
 87. Rahman, M.S.: The advantages and disadvantages of using qualitative and quantitative approaches and methods in language “testing and assessment” research: a literature review. *JEL.* **6**, 102 (2016). <https://doi.org/10.5539/jel.v6n1p102>
 88. Lahrmann, G., Marx, F., Mettler, T., Winter, R., Wortmann, F.: Inductive design of maturity models: applying the rasch algorithm for design science research. In: Jain, H., Sinha, A.P., and Vitharana, P. (eds.) *Service-Oriented Perspectives in Design Science Research*, pp. 176–191. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-20633-7_13
 89. de Bruin, T., Rosemann, M., Freeze, R., Kulkarni, U.: Understanding the main phases of developing a maturity assessment model. In: *ACIS 2005 Proceedings of the American Conference on Information Systems*, p. 11 (2005)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Author Index

A

Abdullai, Larry 478
Abrahamsson, Pekka 173, 190, 231, 299
Adisa, Mikhail O. 478
Agbese, Mamia 231
Ahmad, Noman 265
Ahtee, Tero 190
Akbar, Muhammad Azeem 456
Aldaeej, Abdullah 416
Angarita, Maria Angelica Medina 205
Antonino, Pablo Oliveira 35
Auvinen, Tommi 299
Azad, Nasreen 369

B

Baars, Henning 3
Baninemeh, Elena 327
Birk, Andreas 51
Bjaaland, Ingebjørg Flaata 148
Bjarnason, Elizabeth 360
Bosch, Jan 344
Brink, Henning 493

C

Capilla, Rafael 456
Costa, Inaldo Capistrano 19

D

da Costa, Luiz Alexandre Martins 164
da Silva, Marcelo Augusto 19
Damian, Daniela 132
Das, Teerath 108
Daubaris, Paulius 173
de Oliveira, Fabrício 427
do Outão, Juliana Carvalho Silva 164
Dobslaw, Felix 222
dos Santos, Rodrigo Pereira 35, 164

E

Edison, Henry 360

F

Fronza, Ilenia 173

G

Garidis, Konstantin 283
Ghezzi, Reetta 61, 92
Ghimire, Bachan 132
Guerra, Eduardo Martins 19

H

Halme, Erika 231
Hamza, Muhammad 456
Hannay, Jo E. 148
Haverinen, Henry 400

J

Jansen, Slinger 327
Jormanainen, Ilkka 386
Joutsijoki, Henry 400

K

Kemell, Kai-Kristian 173, 247
Khanna, Dron 416
Knutsen, Leif Z. 148
Koivisto, Miika 108
Kolnes, Martin 205
Korhonen, Minnamaria 92

L

Labunets, Katsiaryna 327
Li, Ze Shi 132
Linkola, Simo 173

M

Mäkitalo, Niko 173
Malcher, Paulo 35
Melegati, Jorge 360
Mikkonen, Tommi 61, 92, 108, 173, 471
Mohanani, Rahul 231, 299
Murray, Alan 283

N

Ngereja, Bertha 148
 Nolte, Alexander 205

O

Öberg, Lena-Maria 222
 Ojikutu, Gbadebo A. 478
 Okker, Timo 299
 Olsson, Helena Holmström 344
 Oyedeji, Shola 478

P

Packmohr, Sven 493
 Päivärinta, Tero 400
 Paloniemi, Teemu 108
 Partanen, Laura 442
 Pattyn, Frédéric 315
 Paul, Fynn-Hendrik 493
 Pekkola, Samuli 77
 Petrik, Dimitri 3
 Piironen, Riina 386
 Porras, Jari 442, 478

R

Rafiq, Usman 315
 Räsänen, Eeli 108
 Rico, Sergio 222
 Rossmann, Alexander 283
 Rousi, Rebekah 173
 Rouvari, Ari 77

S

Sainio, Kari 190
 Samani, Hooman 173
 Serebrenik, Alexander 164
 Setälä, Manu 108
 Sipilä, Antti 442
 Stirbu, Vlad 471

T

Tanilkan, Sinan S. 148
 Toomey, Harold 327
 Tripathi, Nirnaya 265
 Tukiainen, Markku 386

U

Untermann, Anne 3

V

Vakkuri, Ville 173, 247
 Vänskä, Jussi 400
 Viana, Davi 35
 Vilpponen, Hannu 92
 Vuolasto, Jaakko 117

W

Wagenaar, Gerard 327
 Wang, Xiaofeng 315, 427
 Waseem, Muhammad 108

Z

Zaina, Luciana 427