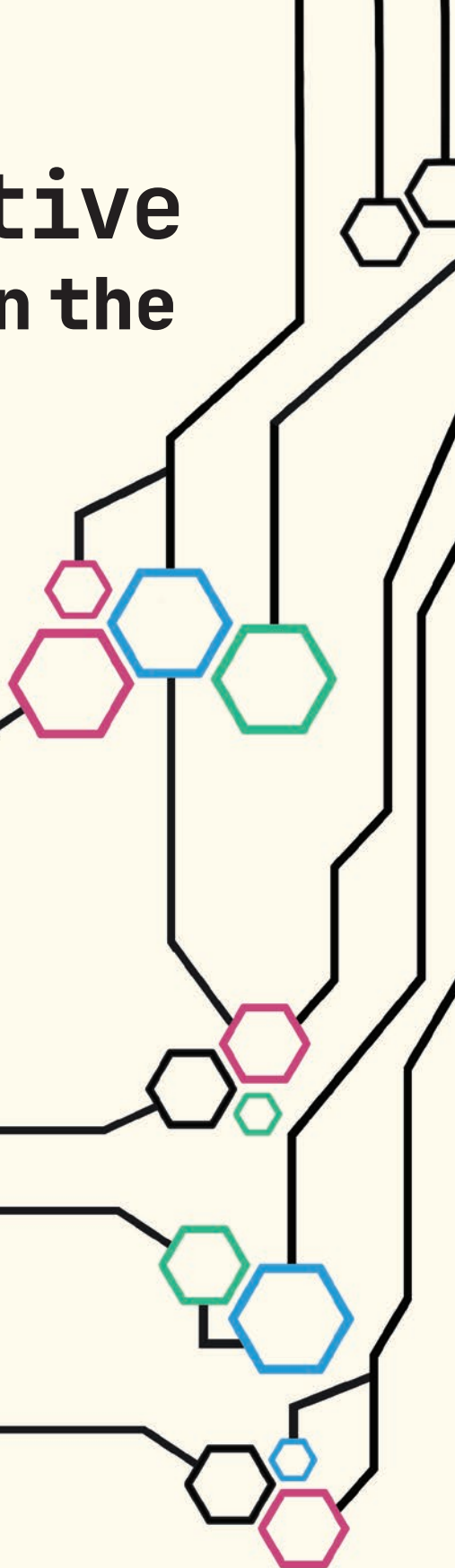


Collaborative Research in the Datafied Society

< Methods & Practices
for Investigation
& Intervention />

Edited by
Mirko Tobias Schäfer
Karin van Es
Tracey P. Lauriault

Amsterdam
University
Press



“Datafication is profoundly changing societies, often in concerning ways. It also opens possibilities for critical social action, so far little explored. This groundbreaking collection convenes a terrific range of leading international writers to showcase collaborative research into this important phenomenon.”
Nick Couldry, Professor of Media, Communications and Social Theory, London School of Economics and Political Science

“Thinking and doing, *mens et manus* this provocative collection of perspectives, methods, theories, and case studies on data and society is above all a testament to the power of collaborative, societally-engaged, and applied research. The plurality of voices and experiences gathered in these pages demonstrate the assertion that the more complex the problem, the more crucial are collaborative solutions.”
William Uricchio, Emeritus Professor of Comparative Media Studies, Massachusetts Institute of Technology

“Schäfer, van Es and Lauriault have brought together an impressive collection of authors, case studies and data practices. Separately and together, the chapters provide theoretical inspiration and practical advice to help the rest of us engage in meaningful collaborations across disciplines and beyond the confines of the university.”
Sally Wyatt, Professor of Digital Cultures, Maastricht University

“A deep dive into dynamics, challenges and urgency of collaborative research in the datafied society, this critical collection showcases how academia and society jointly can drive impactful change. It is an essential reading for anyone willing to engage in research that matters.”
Stefania Milan, Professor of Critical Data Studies, University of Amsterdam and European University Institute

Collaborative Research in the Datafied Society

Collaborative Research in the Datafied Society

Methods and Practices for Investigation and Intervention

*Edited by
Mirko Tobias Schäfer,
Karin van Es, and
Tracey P. Lauriault*

Amsterdam University Press

This work was supported by a Spinoza grant of the Dutch Research Council (NWO), awarded in 2021 to José van Dijck, Professor of Media and Digital Society at Utrecht University.

Cover design: Predrag Markovic

Lay-out: Crius Group, Hulshout

ISBN 978 94 6372 767 9

e-ISBN 978 90 4855 592 5

DOI 10.5117/9789463727679

NUR 133



(<http://creativecommons.org/licenses/by-nc-nd/4.0>)

© All authors / Amsterdam University Press B.V., Amsterdam 2024

Some rights reserved. Without limiting the rights under copyright reserved above, any part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise).

Every effort has been made to obtain permission to use all copyrighted illustrations reproduced in this book. Nonetheless, whosoever believes to have rights to this material is advised to contact the publisher.

Table of Contents

List of Figures	9
Acknowledgements	11
Foreword: The Power of Collaborative Explorations <i>Minna Ruckenstein</i>	13

Part I Theory and Position Papers

1. Making a Difference: The Epistemic Value of Collaborative Research in a Datafied Society <i>Mirko Tobias Schäfer, Karin van Es, and Tracey P. Lauriault</i>	19
2. Performing Critical Data Studies from the Inside: Working with Government to Change Data Regimes <i>Rob Kitchin</i>	35
3. Confronting Politicized Research: The Case for Reflexive Neutrality <i>René König, Payal Arora, and Usha Raman</i>	47
4. Inside Datafication: Entrepreneurial Research for Investigating Emerging Data Practices <i>Mirko Tobias Schäfer</i>	61
5. Open Government Partnership (OGP): Balancing Expertise, Practice, and the Academy <i>Mary Francoli and Daniel J. Paré</i>	87
6. The Challenge of Addressing Subjectivities through Participatory Action Research on Datafication <i>Katherine Reilly and Maria Julia Morales</i>	103

Part II Case Studies

7. Community Responses to Family Violence Policy: A Public Sector Collaboration 121
Anthony McCosker, Jane Farmer, and Arezou Soltani Panah
8. Data Against Femicide: The Process and Impact of Co-designing Digital Research Tools 133
Helena Suárez Val, Catherine D'Ignazio, and Silvana Fumega
9. The Fairwork Project: Promoting Good Labor Practices in the Digital Platform Economy through Action Research 149
Tatiana López, Funda Ustek-Spilda, Patrick Feuerstein, Fabian Ferrari, and Mark Graham
10. Advancing Equity through Data Practices: A Transformative Model for Organizational Change 161
Muna Osman and Hindia Mohamoud
11. Advancing Critical Data Literacy through Justice-Focused Research: A Case Study of the Occupational Hazards of Mass Incarceration 179
Savannah Hunter, Lindsay Poirier, and Nicholas Shapiro
12. Empowering Citizenship Through Academic Practices: The Case Study of Amazonian Civic Media 191
Acilon H. Baptista Cavalcante and Ana Claudia Duarte Cardoso
13. Speculative Data Infrastructures: Prototyping a Public Database on Corporate Tax Avoidance 205
Jonathan W. Y. Gray
14. The DataWorkplace: Collaborative Learning about Datafication in Local Government 219
Krista Ettliger, Mirko Tobias Schäfer, Albert Meijer, and Martiene Branderhorst
15. You Will Be Assimilated: Reflections on Ethnographic Fieldwork on Algorithmic Systems 235
Daan Kolkman

16. Lessons Learned from The eQuality Project: Privacy and Equality for Youth in Networked Spaces <i>Valerie Steeves</i>	251
Afterword <i>Benjamin Peters</i>	261
Bibliography	269
Index	301

List of Figures

Fig. 4.1	Schematic selection process	75
Fig. 8.1	Ideas from co-design sessions for digital tools	139
Fig. 8.2	Data Against Feminicide Highlighter	140
Fig. 8.3	Data Against Feminicide Email Alerts	142
Fig. 10.1	Equity Ottawa theory of transformative change	170
Fig. 12.1	Maps of Terra Firme	192
Fig. 12.2	The nearby periphery districts inside Terra Firme	195
Fig. 12.3	Comparison between Google Maps and Open Street Maps	199
Fig. 12.4	Timeline showing new civic media groups in Terra Firme	201
Fig. 13.1	Table showing missing elements from data proposals	209
Fig. 13.2	Group projects on tax data at King's College London	211
Fig. 13.3	Visualization of uncertainty with the "Fog of Finance" project	213

Acknowledgements

As the many contributions to this volume attest, collaborative research is no easy undertaking. We have, however, been very fortunate to find like-minded students and colleagues to develop research around the impact of digitalization and AI on culture and society. Here, of course, our communities and external partners have been instrumental in their support of these projects and willingness to share their expertise.

For us (Mirko and Karin) the Data School is an exciting place to work. We would like to thank our team: Iris Muis, Jeroen Bakker, Joris Veerbeek, Jill Briggeman, Sofie de Wilde Ligny, Julia Straatman, Alessandra Polimeno, Sander Prins, and Dennis Nguyen; but our gratitude also extends to all the former team members and students whose time with us was productive, inspiring and contributed to what the Data School is today. Thanks for working with us, enabling us to do these projects, keeping us on our toes, and making us better. We are indebted to José van Dijck, who has been supportive of our activities and, through her Spinoza grant, giving us time to be able to develop collaborative research activities. But also, by steering the Governing the Digital Society research area at Utrecht University, she provided us the opportunity to be part of a wonderful and inspiring group of interdisciplinary scholars with whom we collaborate with great pleasure. In this book we claim that universities are changing, and that collaborative research is a promising way forward. This change is not only made possible bottom-up by researchers changing the ways they work, but also top-down by policy makers creating the necessary conditions. We are grateful that Utrecht University is leading the way in this area, and would like to thank them for their trust, for supporting our work, and helping us to succeed: Ted Sanders, Miranda Jansen, Anton Pijpers, Henk Kummeling, and Gerard Barkema.

We would like to thank Tracey P. Lauriault who contributed 3750 Euros from her personal research funds to cover the University of Amsterdam Press Open Access fees, and the MacOdrum Library at Carleton University for contributing \$750 CDN.

A special thanks to Minna Ruckenstein and Christoph Raetzsch for offering critical comments and helpful tips in their reviews. We are very grateful to Jim Gibbons for copyediting the book so swiftly. We would also like to extend our gratitude to all the chapter contributors, most especially all of you beyond the academy who continue to do front line data work.

Lastly, this volume wouldn't have been possible without the support of Chris van der Vegt, a big thank you!

Mirko Tobias Schäfer, Karin van Es, and Tracey Lauriault

Foreword: The Power of Collaborative Explorations

Minna Ruckenstein

This book invites you to think about who you do research with, and what you are trying to accomplish with it. Six years ago, I was gathering evidence in Helsinki for a European mapping exercise of the Berlin-based NGO AlgorithmWatch. The collaboration had started after I received an email asking whether I would be interested in a European-wide mapping exercise of what goes on in the field of automated decision-making (ADM). The aim of this collective undertaking was to get a better sense of emerging cases and debates connected with ADM. We documented locally situated views of how ADM, referring to decisions made without human intervention, was promoted across Europe. The work included national AI strategies and civil society organisations' perspectives, paying attention to regulatory proposals and oversight bodies and mechanisms in place.

The reason I am telling about this collaboration is that it is an example of collaborative research, discussed in this book as joint efforts involving professionals from different backgrounds, disciplines, and organizations to explore a particular topic or project. More importantly, however, this venture made me think about my research in a new way. After the joint report was published, I held workshops with my colleagues where we discussed examples of data analytics, credit scoring and employee evaluations with various audiences, illustrating the risks associated with using machine-generated models in ADM. The discussions in these workshops revealed aspects of automation and algorithmic systems that I could not have learned solely from other researchers.

Opening up research processes to non-academics has been beneficial for my research, but it also benefits the broader society. As this book outlines, educational policies, new funding instruments and commissioned research often expect cooperation that combines the expertise, resources, and

perspectives of researchers and stakeholders to achieve a common goal. On the other hand, the research topic or theme, methodological approaches and aims, might greatly benefit from partnership between academics and stakeholders.

Collaborative research has already proven to be an effective way to study how digital data and algorithmic techniques shape society. Scholars, interested in the implications of datafication and algorithmization, tend to use the term 'algorithmic system' to underscore that it is not merely the algorithm that has sociocultural effects, but the overall system that is defined by people, algorithms, and the organizational or political context. Algorithmic systems are complicated by nature and the information gained is often partial and incomplete. Collaborative research might be an opportunity to observe developing systems, including pilot trials that serve as empirical probes to examine datasets or expectations concerning automation. It can familiarize parties involved in existing data sets, scoring systems and digital tools. By collaborating with practitioners, researchers learn about the visions, practices and methods that are used to promote data uses or algorithmic techniques in organizations. This offers the possibility to pay attention to what kinds of disciplinary backgrounds are favored in teams, what are the most used conversational tropes, tools, and methods, and whether the organization has auditing processes or risk assessment strategies in place.

The volume at hand is a welcome addition at a time when scholars, decision makers and the broader public are increasingly concerned about rapid technological changes, social and political unrest, and ecological disruption – challenges that affect everyone. A focus on collaborative research aligns with the strengthening trend in the humanities and social sciences to move away from scholarly critique towards projects with practical applications and more impactful results. The authors featured in this volume recognize their ability to contribute to relevant societal and policy debates and strive towards more sustainable and just options. This indeed is inspiring. Their efforts highlight the importance of interdisciplinary approaches in addressing complex issues, offering much-needed support for future visions.

The contributions offer timely, reflective and at times provocative accounts of the state of academia, and the possibilities to push forward impactful humanities and social science. On the other hand, the chapters work as a practical guide for students and scholars promoting uses of digital methods and forms of participation to study datafication. With all that is going on, researchers are also proactively transforming higher education, pushing for change, and not merely responding to outside pressure. Some changes are practical, including how to respond to current research funding instruments,

and strengthen infrastructures catering to projects developed in collaboration with external partners. The academic labor required for setting up, managing, and executing collaborative projects is intimately tied to ethical, political, and epistemological stances that matter in terms of how research will be defined and perceived by policy makers and stakeholders in the future.

One of the core features of this book is the thinking and doing that brings people together to share their insights and steers them in collaboration. This also brings to the fore how dialogue with communities and practitioners, whether they are company representatives or civil servants, can only be nurtured in a culture that values openness; ideally, developers of algorithmic systems need to feel that they are participants in an ongoing societal debate. While the authors outline epistemological and political stances that intersect in collaborative research, and offer testimonies of what makes collaboration rewarding, they advocate ‘staying with the trouble’ by paying attention to concerns and contradictions without explaining them away. Their insights underline the importance of treating collaboration as an integral aspect of the research design and knowledge generation. By integrating findings from practical cases that aim to strengthen the study of data-related processes and algorithmic systems, this book illustrates how complex challenges can be more effectively addressed in reflexive dialogue, paving the way for novel approaches in the exploration of data, algorithms, and algorithmic systems.

With this volume, we learn that the gains of collaborative research range from acquiring new insights and accessing data to educating students and professionals. For researchers who aim to contribute to current debates and influence the discourse, collaborative research opens up opportunities to develop participatory, adaptive, improvisational, and inventive methodologies. This volume celebrates the agentic capacity of researchers to create their own paths through collaboration. To keep this exploratory spirit alive, it is crucial to continue fostering cross-organizational reflection and dialogue – a goal this book will support successfully.

About the Author

Minna Ruckenstein is professor of Emerging Technologies in Society at the University of Helsinki and the founder of the Datafied Life Collaboratory. She has dedicated over a decade to studying the human aspects of digitalisation, datafication, and AI. Currently funded research projects focus on public values in algorithmic futures, re-humanising automated decision-making, and repair and renewal of algorithmic systems.

Part I

Theory and Position Papers

1. Making a Difference: The Epistemic Value of Collaborative Research in a Datafied Society

Mirko Tobias Schäfer, Karin van Es, and Tracey P. Lauriault

Abstract

This chapter addresses the evolving role of academia amidst budget constraints and neoliberal policies, highlighting the growing need for its work to be more socially relevant, especially in the humanities. It argues that academia can actually benefit from moving beyond its institutional walls, engaging with diverse community and civil society stakeholders. Such collaboration enables universities to respond to pressing societal challenges. The chapter explores three primary motivations for increased academic engagement with societal sectors, identified by researchers and university administrators: vocational, educational, and societal impetus, and advocates for a fourth motivation: the epistemic impetus. Collaborative research allows researchers to gather evidence and generate insights to produce knowledge with communities and in context, enriching academic research and allowing interventions and the application of findings.

Keywords: Societal engagement; Data work; Stakeholders; Civil society; Unacknowledged labor

Academia is rapidly changing, struggling with relevance and the need to be more engaged with the community. Also, universities are experiencing significant budget cuts. While austerity measures and neoliberal ideologies have sparked debates about the value of teaching and research, particularly in the humanities, for some time (e.g., Brandt 2011; Collini 2012), university policies increasingly push for greater engagement with social sectors. We

consider this shift to be positive and timely, given urgent global challenges that research institutions (among others) must address: the climate crisis, pandemics, hybrid warfare, migration, demographic shifts, digitalization including AI, and the platformization of social services to name a few. Here, scientific research and social, technical, and political discourses and public debates align with those of journalistic commentary and advocacy from political groups, industry, and civil society. Reflecting on these intersections, we argue that academia has much to gain by moving beyond its institutional wall, in our case, by doing data work with a community.

The contributors to this book share a specific interest about datafication, knowledge economies and the rise of artificial intelligence (AI). They collaborate with stakeholders across diverse communities and civil society to tackle challenges that address pressing issues stemming from data practices and social justice issues. The chapters also discuss public engagement that extends beyond the scope of traditional science communication, teaching, and the reporting of applied research results. They also constitute transdisciplinary cooperation and mutually generated knowledge with actors outside the university, data work with issue stakeholders, professionals, practitioners, the public and researchers from other multiple disciplines.

Reflecting upon this contemporary context, we identified three main arguments put forth by researchers and university administrators for more collaboration with societal sectors, and we emphasize a fourth:

1. Our expertise is needed (vocational impetus).
2. It provides learning opportunities (educational impetus).
3. We can contribute to shaping society (societal impetus).

The *vocational impetus* acknowledges that research produces knowledge to respond to societal challenges, and for researchers to make efforts to mobilize their findings with specific audiences, and to engage with stakeholders who can build on that expertise. For emerging data and AI practices, this increasingly includes researchers from the humanities and the social sciences (Van Dijck 2017).

The *educational impetus* is about creating learning opportunities resulting from cooperation between universities and civil society. Here, universities find new student audiences, develop community-engaged learning, and provide education for professionals (e.g., UNESCO 2023). The *societal impetus* targets the application of research findings for the benefit of society at large. Utrecht University and Carleton University, among others, have both expressed these sentiments in their mission statements: “Utrecht University

is working to create a better world” (Utrecht University 2020) and “connect and partner with community and businesses to advance excellence in research” (Carleton University 2024).

However valid such motivations are for encouraging and intensifying collaboration between academia and society may be, they typically imply the view that knowledge originates within academia and is then distributed to society as a whole. We, however, propose an additional motivation: the *epistemic impetus*. This epistemic impetus extends the capacity to gather evidence and arrives at insights to produce knowledge with community and in context.

Datafication, we suggest, also provides researchers and community with new data resources, including advancements in computational or digital methodologies, allowing for new possibilities to generate and access new kinds of empirical evidence which had previously been out of reach (Van Es and Schäfer 2017). New data and computational approaches necessitate different kinds of interdisciplinary collaboration, to jointly refine knowledge about the impact of data on our lives, work, politics, and culture. Collaborative research that actively engages with society not only taps into the intrinsic motivation for knowledge discovery – the *epistemic impetus* – but also increases opportunities for professional and educational growth which, in turn, supports a broader commitment to the *societal impetus*.

Doing Data Work

The emerging field of critical data and AI studies has provided a much-needed correction with regard to the objectivity claims made in the overly optimistic discourse of Big Data (e.g., boyd and Crawford 2011; Iliadis and Russo 2016; Richterich 2018; Kitchin and Lauriault 2022) and is aligned with a similar critique of AI (e.g., Broussard 2018; Crawford 2021; Buolamwini 2023). Collaborative research in this domain not only facilitates the collection, presentation, and discussion of empirical evidence to support this critique but also actively enables education, intervention, and transformative change in data, AI, and social-technical processes.

Collaborative research practices, involving colleagues from various disciplines with practitioners and issue specific stakeholders, aim to explore opportunities for the in-depth study of social and technological transformation. Collaborating with stakeholders and the co-creation of research leads to insights beyond the typical distant gaze of traditional academic research. This form of shared inquiry also addresses issues, problems, and challenges

often overlooked and willfully ignored. This may involve reexamining labor practices and organizational structures, governance, developing data literacy and the co-creation of datasets to name a few. This is what we call *data work*.

Data work includes meticulous and labor-intensive efforts along the data value chain: capturing and annotating data; compiling datasets; storing, retrieving, and processing data; as well as visualizing and publishing data. Additionally, the concept accounts for the social and local contexts within which data practices occur. Data work here is about the detailed, socially conscious effort to counter the often unsubstantiated promises associated with big data. Collaborative research projects in this area facilitate and advance data work efforts, enabling effective knowledge transfer and interventions. They also benefit the research methods, as researchers develop the capacity to count, measure, and record events or phenomena previously not available as data. Unlike easily available records, which may be of lesser quality in representing phenomena but make up what has widely been called “Big Data,” this kind of data work also represents “real social analytics” (Couldry et al. 2016). Furthermore, the close study of a phenomenon enables critical examination of the indicators used to score individuals, a practice central to many reported cases of algorithmic harm. Collaborative research with knowledge experts beyond the academy brings us closer to the actual phenomenon, putting all in a better position to engage in “data point critique” (Gerlitz 2017). This is connected research that responds to societal needs and urgencies, most especially datafication and algorithmic issues that affect local, social, and organizational data and technological contexts (e.g., Loukissas 2019).

Shifting Academia

Universities are struggling for social relevance, to be more open and inclusive, and wanting to be oriented toward addressing societal issues. We also see scholars and administrators increasingly participate in public debates, influence policy decisions, and create educational outreach programs. Further, there is also the move toward the mobilization of scholarly knowledge beyond the academy. This has long been subject to criticism, in some cases critiques that universities have become more commercial and are limiting academic freedom, stifling methodological rigour, and cutting costs (e.g., Slaughter and Rhoades 2000; Bok 2003; Oostrom 2007). Indeed, applied research and life-long learning have become the hallmarks of neoliberal

policy to serve the research-and-development (R&D) and the human-capital demands of corporations. The aim of demonstrating value to society has fuelled new managerialist outlooks, such as impact agendas (Holbrook 2017). We and the authors of the chapters in this book, on the other hand, argue for a different type of impact and engagement, by providing examples that make a difference in the datafied society. Approaches discussed in this volume not only facilitate applied data research but also enrich fundamental research by providing new insights and perspectives. We embrace the power of collaborative, societally engaged, and impact-focused research as part of the *epistemic impetus* that has social and technical relevance.

Through collaborative research, universities can advance the *educational impetus* as they involve students in research activities, who often become professionals who continue cooperating with researchers, and actively inform public policy and provide applicable real-world solutions to the datafied society (Lauriault, Leonne and Ivanoff 2021; Schäfer, Van Es, and Muis 2023). In other words, doing data work helps develop highly qualified personnel for social good – the *vocational impetus*. We align with the recently broadened understanding of open science as being research practices beyond open data and open publishing (Miedema 2022). Through enlarging the field of data work with a holistic understanding of open knowledge production processes with stakeholders from societal sectors and with those experiencing social and economic exclusion or societal harms, our research activities become meaningful and engaged data work.

Collaborative research considers the expertise found among stakeholders and their respective societal sectors and communities. Changing research practices, as described here, allow researchers to immerse themselves in specific societal sectors and communities under investigation and to study phenomena up close with local experts. The practices discussed here extend beyond participatory observation: they are forms of action research, where scholars, practitioners, and the public shape the research questions and methods and do data work together to address immediate issues, achieving results that could only be realized in multi-sectoral teams (Cizek and Uricchio 2022; Dwivedi 2024). Collaborative research here is rooted in disciplinary training, but it unfolds in interdisciplinary and even transdisciplinary research projects; it takes different manifestations of expertise into consideration and is driven more by the current urgencies and needs felt in particular societal sectors than by the inner, often self-referential discourses of academic sub-disciplines and associations. This approach represents a significant shift from focusing on highly individualized academic achievements to valuing collective work.

Researchers from different disciplines collaborate with practitioners or stakeholders to conduct a research project. In such collaborations, the distinctions between participants outside of the university, academic researchers, and support staff often blur in favor of a cooperative effort. These are historically informed and situated perspectives of cultural complexity that are socially dynamic, involve public discourses and aesthetics, and require the integration of the humanities and the social sciences (Van Es and Schäfer 2017). This also means that research will develop directly with events unfolding within particular societal or community contexts. While we do not argue that “pure science” will or should be replaced, we do however think that this kind of data work invites humanities researchers, social scientists, and data scientists to alter their modes of research.

In *Changing Cultures in Higher Education*, Tony Bates (2010, 22) quotes a vice chancellor’s metaphor to describe universities’ resistance to change: “Universities are like graveyards. When you want to move them, you don’t get a lot of help from those inside.” Such resistance is not limited to “pure scientists” who would prefer to continue not to reward, or even acknowledge, individual academic labor being performed outside the norms of the peer-reviewed publication or grants. When universities decide to collaborate with partners outside the university, and have their researchers engage with publics outside of the reigning academic discourse to educate audiences other than the traditional student cohorts, data work profoundly changes. These changes are as much epistemological as they are administrative challenges.

As interdisciplinary and even transdisciplinary research practices become more relevant, academic institutions should reconsider how initial disciplinary training and the subsequent inter- and transdisciplinary collaboration will be ensured. This shift requires the additional education of researchers in neighboring disciplines, the development of skills for working in mixed teams, and the need to listen and understand different perspectives. It raises questions about the rather monolithic organization of universities into departments defined by disciplines, a structure that often seems to stifle collaboration between disciplines rather than stimulating it.

With regard to operations, universities need to expand their research support offices. Up to now they have catered to the traditional research grants distributed through national or supranational funding organizations or various foundations. Legal advice, expertise in contracting, project management, data sharing agreements, progressive data and technology procurement, and corporate communication are usually underdeveloped in supporting collaborative research projects with stakeholders or funders

other than with traditional funding organizations. If such knowledge is present, it mostly relates to joint projects with large corporations. As many of the chapters in this book demonstrate, incredibly relevant work is done with small and local organizations and communities.

Aside from the absence of suitable policies, engagement guidelines, and essential support infrastructure, the primary barrier to collaborative research is the lack of recognition and rewards given to do this sort of work. University administrations may claim the contrary, but traditional incentives remain prevalent within academic institutions. These incentives primarily benefit the individual researcher, typically a full professor supported by a postdoctoral fellow and several PhD students. In evaluations and considerations for potential promotions, the metrics that carry weight include the publication of peer-reviewed papers, the acquisition of grants, the quantity of supervised and successfully defended dissertations, and, occasionally, public visibility through television appearances or op-ed contributions in newspapers. There are very few formal incentives for researchers to go out of their way to build a network with partners outside the university, to engage in the challenging process of interdisciplinary collaboration, to meticulously review results for practical problem-solving and application, to involve their students in field research experiences, or to develop innovative educational formats for professionals encountered in the field and with government.

This book identifies various challenges in the domain of collaborative research. Despite these obstacles, as the examples presented here demonstrate, many have successfully engaged in collaborative research, gaining insights they could not have captured otherwise, intervening effectively in society, providing much-needed data work, and enabling others through it, and taking part in shaping the digital society.

Our aim with this contribution is to facilitate connections among colleagues across different disciplines who are already involved, in some capacity, in similar research efforts. We hope to inspire others who are addressing issues relevant to current societal challenges, and those compelled to interact with stakeholders and audiences beyond the academic sphere, by providing them with practical guidance on practices and methods for setting up collaborative research projects, and to lead scholars in new directions. Finally, we want to inform university policymakers about the challenges posed by collaborative research, encouraging them to design policies that support these endeavors and appropriately recognize and reward the efforts and accomplishments of university faculty and employees involved in such work.

Collaborative research fosters practices, strategies, and tactics that effectively address the challenges of doing this type of research. These approaches can inform university policies and shape the training of scholars to respond appropriately to these challenges. We suggest that the collaborative research not only exemplify the four impetuses – vocational, educational, social, and epistemic but also fall into two broad objectives:

1. Making visible and acknowledging forms of (academic) labor, valid research, and teaching efforts which are widely marginalized.
2. Highlighting exemplary practices, activities, and methodologies that respond to the shift of academia towards intensified societal engagement.

This book provides insight for universities to be more closely linked with public debates, societal needs, and pressing issues in spite of the fact the lack of their institutional infrastructure to effectively support this type of public engagement and collaborative research.

Overview of This Book

This book provides examples of collaborative research that address the consequences of datafication, and excellent examples of doing data work. These research activities are inherently interdisciplinary and multi-sectoral that includes the expertise of practitioners, stakeholders, or the public. Chapters include practices that respond to urgencies within the social sector or communities by researchers and facilitate mutual knowledge transfer. They also take into account situated knowledge and practices rather than prescriptive analysis from afar, as here research outcomes are co-created with stakeholders.

These efforts align with the emerging ambitions of universities to become more inclusive, acknowledge diversity, expand their curricula, and develop community-engaged learning. The activities described incorporate such practices naturally. The contributors to this book are developing new forms of data informed interdisciplinary inquiry and are re-shaping the discourse of datafication and doing data work.

The book is divided into two parts. Part I includes theoretical perspectives and position statements towards the development of evidence-based and impactful collaborative research from different universities and in diverse societal contexts. They emphasize cooperation with external partners and how these processes affect and shape their research. Drawing from their

experience, they address the limitations imposed by how academic research has typically been organized and point to necessary changes to normalize these activities in academic institutions.

Part II presents ten case studies. Here we selected novel data work practices that engage with stakeholders, collaborate with partners outside the university, and use knowledge resources beyond traditional institutions. More importantly, we sought insightful discussion about the tactics, resources, and skills necessary to conduct this type of datafied society research. Many of these case studies demonstrate a commitment to improve the situation of people in the contemporary digital society, to develop novel didactic knowledge transfer processes, and, most importantly, to apply academic labor to societal sectors. This includes case studies that provide clear evidence of societal impact, characterized by successful knowledge mobilization and collaboration with partners outside the university. We hope these case studies inspire others and serve as models to establish new forms of data work research initiatives.

Theoretical Perspective and Position Statements

Part I begins with two chapters that address roles where scholars are neutral observers or active agents of change and discuss questions pertaining to academic neutrality and independence. In the chapter “Performing Critical Data Studies from the Inside: Working with Government to Change Data Regimes,” Rob Kitchin revisits the work he and his team have done cooperatively with government to change data regimes. He argues that academics should operate beyond the boundaries of their individual disciplinary contexts and actively engage with relevant problems. Drawing from almost two decades of research, the chapter demonstrates that active engagement in society and working with those who have an insider position is often the most efficient way to develop relevant policy, shape government programs, and build mutually beneficial infrastructure. In their chapter “Confronting Politicized Research: The Case for Reflexive Neutrality,” René König, Payal Arora, and Usha Raman provide a way forward with the concept of reflexive neutrality that strikes a balance between the so-called “neutral” researcher and the counterproductive activist researcher. Referring to Pielke’s notion of the researcher as an honest broker of different policy options, they propose reflexive neutrality as a process by which researchers acknowledge their sociopolitical embeddedness while formulating evidence-based results from their analysis.

The next set of chapters reflect on applied research, methods of collaboration, and the relationship between research, teaching, and society. In his chapter, Mirko Tobias Schäfer introduces entrepreneurial research as a distinct method to investigate data practices and AI in their social contexts. Using entrepreneurial activities, he argues, allows researchers to immerse themselves in specific societal sectors to arrive at privileged insights to create learning opportunities and possibilities for intervention. Referring to examples from the work done at the Data School at Utrecht University, Schäfer points to the challenges and opportunities emerging from such an approach. Subsequently, in “Open Government Partnership: Balancing Expertise, Practice, and the Academy,” Mary Francoli and Daniel J. Paré explore the professional benefits and the challenges they encountered in their data-focused policy work for, and engagements with, the Open Government Partnership. They argue that applied research nurtures and enriches service, scholarship, and teaching, yet they find that the value and merit of such work is often not recognized in academia. Their chapter formulates concrete measures at institutional and individual levels. In their chapter “The Challenge of Addressing Subjectivities through Participatory Action Research on Datafication,” Katherine Reilly and Maria Julia Morales write about data audits as a mode of participatory action research. Drawing from various research projects, they argue that general data literacies fall short, and suggest that making a meaningful difference requires active participation in reviewing and situating data within affected communities. Their chapter also connects critical data studies to participatory action research to demonstrate how the involvement of researchers and citizens in actual data projects might lead to improvements in data literacy, citizen agency, and research quality.

Case Studies: Do Try This at Home!

In part II authors provide ten case studies clustered around three themes: (1) accountability and policy work; (2) data work and literacy; and (3) collaborative practices. Part II starts with cases concerning the building of infrastructure for accountability and informing policy. “Community Responses to Family Violence Policy” describes the collaboration between public management employees and researchers to chart how public awareness and the understanding of family violence has changed over time in response to public interventions and policy using novel data analysis techniques. Anthony McCosker, Jane Farmer, and Arezou Soltani Panah also reflect on

the quality of the collaboration, the motivation behind their joint research approach, and its outcomes.

“Data Against Femicide” provides a detailed account of the process and impact of feminist participatory approaches that place technology development and data science in the service of activists and social movements. In their mixed team of activists and researchers, Helena Suárez Val, Catherine D’Ignazio, and Silvana Fumega developed practical tools to collect evidence about femicide from media reports in Latin America. These tools made femicide visible and became a call to action to authorities, who had been neglecting available evidence and also ignored an entire category of criminal activity.

“The Fairwork Project: Promoting Good Labor Practices in the Digital Platform Economy through Action Research,” by Tatiana López, Funda Ustek Spilda, Patrick Feuerstein, Fabian Ferrari, and Mark Graham, introduces a method to tackle information asymmetries in the platform economy that construct and conceal exploitative labor relations. The authors provide critical insights into the challenges of conducting action research in the gig economy with regard to maintaining the independence of the research process and its findings. In “Advancing Equity through Data Practices” from Equity Ottawa, a program working to advance the integration of immigrants in Ottawa, Canada across a multi-sectoral partnership, Muna Osman and Hindia Mohamoud explain how the project supported processes of organizational change at the intersection of equity and the datafication of public institutions. This is data work that also reflects upon successes, challenges, and next steps.

Attention then shifts to cases that empower through data literacy. In “Advancing Critical Data Literacy through Justice-Focused Research,” Savannah Hunter, Lindsay Poirier, and Nicholas Shapiro report on a project that uses open government data to explore Occupational Safety and Health Administration violations in US prisons and detention centers. They reflect on how the project, involving various faculty members and (under)graduates, advanced critical data literacy education for students, and they consider the challenges involved in this sort of multidisciplinary collaboration.

Acilon H. Baptista Cavalcante and Ana Claudia Duarte Cardoso then offer insights from the Data Firme project in their chapter “Empowering Citizenship through Academic Practices.” Using Design Thinking, they discuss how academics co-created media practices with young residents from Terra Firme, a neighborhood in the Brazilian city of Belém, to counter representations of that area in the mainstream media and to strengthen a civic media network. Their contribution reflects on the academic challenges of a project built in collaboration with the community.

Jonathan Gray, in “Speculative Data Infrastructures: Prototyping a Public Database on Corporate Tax Avoidance,” explores three ways to workshop “data-in-the-making.” These workshops critically engage with “Country-by-Country Reporting” (CBCR) data. In organizing collaborations with data and offering a generative format, they promote collective learning and interpretation among researchers, teachers, students, and activists. In what follows, “The DataWorkplace: Collaborative Learning about Datafication in Local Government” offers insight into data work between a university and local and regional governments around how these organizations adapt to data practices and AI. Krista Ettliger, Mirko Tobias Schäfer, Albert Meijer and Martiene Branderhorst then reflect on transdisciplinary research as a means to tackle such complex problems. Next, in “You Will Be Assimilated,” Daan Kolkman discusses his own experiences conducting ethnographic fieldwork about data professionals and the algorithmic systems they helped create. He reflects on navigating the challenges to negotiate access, establish rapport, and develop expertise. Additionally, he considers the implications of his gradual transition from an outsider to an insider’s role regarding research integrity and established academic practices of rewards and recognition. Finally, in “Lessons Learned from the eQuality Project,” Valerie Steeves reports on an interdisciplinary and intersectoral project concerned with creating knowledge about young people’s lived experiences of privacy/surveillance and equality in networked space. She discusses the benefits and challenges encountered in a partnership approach to research.

Taken together, the book’s chapters formulate relevant concepts for grounding societally engaged research in the theories and methodologies from different disciplines. Their authors also redefine what is commonly understood as academic research, and they make unacknowledged academic labor and data work explicit. They advocate for urgent changes to be made to the traditional organization of universities, marked by the latter’s counterproductive distinction between research and support staff; insufficient support infrastructures for community-engaged learning and the co-creation of research with stakeholders and partners outside the university; their neglect of societal impact work; and the hesitancy of their commitment to inter- and transdisciplinary collaboration.

In the afterward, Ben Peters considers practice as theory that matters. He strongly argues that we revisit our research priorities and embrace collaborative research as a means forward so that the inherent drive of fact-finding and world-understanding can be combined with developing capacities for actually making a difference.

In summation, we seek, with this book, to highlight the unique knowledge gained through collaborative research, which provides empirical insights that may be challenging or impossible to achieve through other methods. Besides generating practical outcomes, educational frameworks, and sustainable knowledge exchange between academia and society more broadly, collaborative research emerges here as a highly beneficial and essential practice for our knowledge-driven and datafied societies!

References

- Bates, Tony. 2010. "New Challenges for Universities: Why They Must Change." In *Changing Cultures in Higher Education: Moving Ahead to Future Learning*, edited by Ulf-Daniel Ehlers and Dirk Schneckenberg, 15–25. Berlin: Springer.
- Bok, Derek. 2003. *Universities in the Marketplace: The Commercialization of Higher Education*. Princeton, NJ: Princeton University Press.
- boyd, danah, and Kate Crawford. 2011. "Six Provocations for Big Data." *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*.
- Brandt, Reinhard. 2011. *Wozu noch Universitäten?: Ein Essay*. Hamburg: Felix Meiner Verlag.
- Broussard, Meredith. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.
- Buolamwini, Joy. 2023. *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*. New York, NY: Random House.
- Cizek, Katerina and William Uricchio. 2022. *Collective Wisdom: Co-Creating Media for Equity and Justice*. Cambridge, MA: MIT Press.
- Collini, Stefan. 2012. *What Are Universities For?* Penguin UK.
- Couldry, Nick, Aristeia Fotopoulou, and Luke Dickens. 2016. "Real Social Analytics: A Contribution Towards a Phenomenology of a Digital World." *British Journal of Sociology* 67 (1): 118–137.
- Crawford, Kate. 2021. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.
- Dwivedi, Yogesh K., et al. 2024. "Real Impact: Challenges and Opportunities in Bridging the Gap Between Research and Practice—Making a Difference in Industry, Policy, and Society." *International Journal of Information Management* 78: 102750.
- Holbrook, J. Britt. 2017. "The Future of the Impact Agenda Depends on the Revaluation of Academic Freedom." *Palgrave Communications* 3 (39): 1–9. <https://doi.org/10.1057/s41599-017-0041-0>.

- Iliadis, Andrew, and Federica Russo. 2016. "Critical Data Studies: An Introduction." *Big Data & Society* 3 (2): 2053951716674238.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Miedema, Frank. 2022. *Open Science: The Very Idea*. Dordrecht: Springer.
- Oostrom, Frits. 2007. "Markt en ziel. De universiteit als instituut van de verbeelding." *Ons Erfdeel* 50 (4): 54–65.
- Richterich, Annika. 2018. *The Big Data Agenda: Data Ethics and Critical Data Studies*. London: University of Westminster Press.
- Schäfer, Mirko Tobias, Karin van Es, and Iris Muis. 2023. "Investigating the Datafied Society: Entrepreneurial Research as an Approach." In *Situating Data: Inquiries in Algorithmic Culture*, edited by Karin van Es and Nanna Verhoeff, 267–272. Amsterdam: Amsterdam University Press.
- Slaughter, Sheila, and Gary Rhoades. 2000. "The Neo-liberal University." *New Labor Forum* 6: 73–79.
- Smith, Simon, Vicky Ward, and Allan House. 2011. "'Impact' in the Proposals for the UK's Research Excellence Framework: Shifting the Boundaries of Academic Autonomy." *Research Policy* 40 (10): 1369–1379.
- UNESCO Institute for Lifelong Learning, Shanghai Open University. 2023. International Trends of Lifelong Learning in Higher Education. Research report. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000385339>.
- Utrecht University. 2020. *Strategic Plan 2025*. <https://www.uu.nl/en/organisation/strategic-plan-2025/the-strategic-plan>.
- Van der Zwaan, Bert. 2017. *Higher Education in 2040: A Global Approach*. Amsterdam: Amsterdam University Press.
- Van Es, Karin, and Mirko Tobias Schäfer. 2017. *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.

About the Authors

Mirko Tobias Schäfer is Associate Professor of AI, Data & Society at Utrecht University's research area *Governing the Digital Society* and the Department for Information & Computing Sciences. Mirko is co-founder and Sciences lead of the Data School. He studies the datafication of public management and engages in the development of responsible and accountable AI and data practices.

Karin van Es is Associate Professor, Media and Culture Studies, and project lead Humanities at Data School at Utrecht University. Her research

contributes to the fields of software studies and critical data studies. She currently studies streaming video platforms, examining media policies and exploring innovative methods to analyze these services.

Tracey P. Lauriault is Associate Professor, Critical Media and Big Data and board member of the Institute for Data Science at Carleton University in Canada. As a data and technological citizen, she examines data and technological systems to make them more just, inclusive, equitable and environmentally sustainable.

2. Performing Critical Data Studies from the Inside: Working with Government to Change Data Regimes

Rob Kitchin

Abstract

This chapter considers the role of academics in society and to what extent they should seek not simply to produce knowledge about the world but to change it. The chapter argues that academics should operate beyond the academy, proactively engaging state, industry, and civic society organizations to enact progressive interventions. It contends that the most effective way to achieve such interventions is to occupy insider positions that directly contribute to the formulation of policy and the development of programs and infrastructures. The case is made by reflecting on nearly two decades of applied action research, working with state agencies and government departments to try to change data policies and practices and to build data infrastructures.

Keywords: Critique; Advocacy; Action research; Positionality

Introduction

There is a long-standing debate in the social sciences concerning the relevance, purpose and practice of academic research (Mouton and Marais 1988; Fuller and Kitchin 2004a; Bastow et al. 2014). The dynamics in these debates largely hinge on beliefs regarding the extent to which: (i) research should simply produce knowledge about the world or, alternatively, should actively seek to change it; and (ii) academia should work with and for state and industry actors. Some would hold that the academy should produce

impartial and independent knowledge, with researchers' personal politics left at the university gate as they conduct objective and neutral research and act in a value-free manner. In this view, academia is to produce independent, objective, and impartial assessments of society and to set out the pros and cons of policy options that enable others to evaluate such assessments and make their own decisions (Mouton and Marais 1988; Kitchin 2015). In other words, scholars produce knowledge and suggest possible ways to respond to their findings, but it falls to others to evaluate and apply such knowledge in practice. This position does not preclude conducting commissioned research on behalf of state and industry actors; however, such research should be objective, impartial, and non-prescriptive.

In contrast, the post-positivist social sciences hold that it is impossible to produce truly independent and impartial knowledge; the work of an academic, and of the academy more broadly, is always political (Rose 1997; Fischer 2005). These inherent politics, rather than being denied, are mobilized within academic endeavor (research, teaching, publishing, external service) to actively seek change in society. For some, change is sought principally through critique, which assesses and appraises the work of the state and of industry, which are held accountable for their expressions of power and structural violence (Mitchell 2004; Allen 2011). Here, the academy is necessarily separate from the state and industry (Allen 2011), ensuring its critical distance, independence, and scientific autonomy rather than being co-opted into and legitimizing state and industry actions (Wilton 2004; Allen 2011). For others, critique is not a sufficient intervention in the process of seeking change. Those holding this view propose that academics work with civil society organizations and local communities, performing action research, formulating and promoting policy alternatives, and undertaking advocacy and activism (Fuller and Kitchin 2004b). Such activity can include direct work with state and industry actors, even though they may hold different views and aspirations, seeking change from within. The latter approach has become more popular in recent years given the prevalent impact and engagement agenda, along with the pressure being exerted by government policy and by funding agencies demanding that the academy produce work with instrumental value (Bastow et al. 2014).

This variance in the understanding of the politics, praxes, and purposes of academic endeavor is partially encapsulated in Michael Burawoy's (2005) taxonomy of social science research. Burawoy argues that there are four main orientations and praxes of social science, which are defined principally through the form of knowledge produced (instrumental or reflexive) and

the intended audience (academic or extra-academic). Independent research, of the supposedly value-free and neutral variety, as well as the sort that engages in critique and advocates for change, is largely subsumed under instrumental and reflexive knowledge. In contrast, research produced for an extra-academic audience is more applied, engaged, and action-oriented. Of course, the production of reflexive knowledge can be translated into instrumental insights, and work produced for academic peers can be recast to suit extra-academic audiences, and vice versa. An academic can certainly undertake research that fits each category: that is, it is wholly possible to produce instrumental and reflexive knowledge for academic and extra-academic audiences. For example, over the course of my career, I have produced instrumental knowledge for an academic audience, but I have also conducted applied and policy work for an extra-academic audience, produced reflexive knowledge that is highly critical of the state and industry, worked on participatory projects with community groups, and produced public scholarship (blogs, social media, print and online media) aimed at influencing public opinion.

Table 2.1. Forms of knowledge production and praxis (framed with respect to Geography)

	Academic audience	Extra-academic audience
Instrumental knowledge	<i>Professional Geography</i>	<i>Applied (Policy) Geography</i>
– Knowledge	Theoretical/empirical	Concrete
– Legitimacy	Scientific norms	Effectiveness
– Accountability	Peers	Clients/patrons
– Pathology	Self-referentiality	Servility
– Politics	Professional self-interest	Policy intervention
Reflexive knowledge	<i>Critical Geography</i>	<i>Participatory and Public Geography</i>
– Knowledge	Foundational	Communicative
– Legitimacy	Moral vision	Relevance
– Accountability	Critical intellectuals	Designated publics
– Pathology	Dogmatism	Faddishness
– Politics	Internal debate	Public dialogues

Source: Adapted for Geography by Kitchin et al. (2013) from Burawoy (2005)

Critical Data Studies (CDS) makes data and its framing, production, and use its core conceptual, analytical and empirical focus (Dalton and Thatcher 2014). The “critical” aspect in its name refers to the adoption of a perspective rooted in critical social theory. CDS does not regard data or data work as commonsensical or take them at face value (Kitchin 2022).

Rather than data being understood in essentialist terms (natural, benign, representative), abstracted from the world in neutral and objective ways subject to technical constraints, data are understood to be *produced*; that is, they are generated within and reflective of socio-technical contexts. Data are not simply waiting to be collected; they do not exist before they are generated (Markham 2017). Data are the product of discursively framed and technically mediated processes as shaped by protocols, organizational processes, measurement scales, categories, and standards that are designed, negotiated, and debated (Kitchin 2022). Similarly, the entirety of the data lifecycle (generation, handling, processing, storing, sharing, analysis, interpretation, deletion) is socio-technically mediated and saturated with politics. So, too, are the production and operation of data infrastructures and the many ways in which data are used (Leonelli et al. 2017). CDS, then, reflects on philosophical concerns relating to data and their use, and it asks political and ethical questions so as to reveal what is really at stake in data-driven systems and regimes (Kitchin 2022). As such, the majority of endeavors in CDS research are reflexive and academic (critical), though a reasonable proportion of them are instrumental and extra-academic (applied), and some are reflective and extra-academic (participatory). In contrast, research in Data Science is dominated by instrumental and academic (professional) and instrumental and extra-academic (applied) endeavors.

My own CDS research has predominantly been critical and applied rather than professional or participatory/public. I would consider myself a post-positivist scholar who seeks to change the world in proactive ways. My work is political and pragmatic; it engages and works with public and civic stakeholders and is aimed at multiple audiences. Consistently critical, it is also applied work, seeking to formulate and shape public policy and practice and to help build public data infrastructure. At times, it is participatory and activist or takes the form of advocacy and public debate. Over the past twenty years, I have enacted this multifaceted approach with respect to state data and associated data assemblages and infrastructures in Ireland. In what follows, I focus on work relating to planning and development, discussing three projects (All-Island Research Observatory, the Programmable City project, and the Building City Dashboards project). I have undertaken similar work related to culture and heritage as a co-principal investigator for the Digital Repository of Ireland, as well as to the qualitative social sciences as a co-principal investigator of the Irish Qualitative Data Archive and as a participant in the open data movement more generally.

Performing Critical Data Studies from the Inside

My research related to data is premised on the assumption that the best way to enact positive progressive change is not to produce knowledge and leave it to others to convert such knowledge into necessary action, but rather to seek the desired outcome proactively. However important critical analysis for a mainly academic audience may be in terms of generating fundamental insights and informing praxis, its impact beyond the academy is usually limited if it is not accompanied by translational praxis work. If one really wants to influence how the state deals with a phenomenon, the optimal approach is to directly contribute knowledge, ideas, potential solutions, and resources (e.g., time, energy, networks, institutional capacities, reputation, etc.). If one regards a new policy to be necessary, then the way to push it onto an agenda is to be an advocate and to lobby for its creation. If one desires a particular policy formulation, it is best achieved by being involved in its creation. The mere existence of a policy does not mean that all will welcome and implement it, so if one wants to ensure its adoption and a desired impact, then advocacy work needs to be undertaken to persuade others to engage and implement its core principles and practices. If one believes that data practices and management need to be updated or rethought, then offering constructive feedback on better alternatives, or facilitating scoping workshops, or delivering training for staff, will help shift embedded thinking and processes in new directions. If one wants to support the creation of a data infrastructure and seeks it to possess certain qualities and capacities, then the best means of ensuring its production is to be part of the development team. In other words, the best way to effect the desired change is to gain an insider position.

A cynic might argue, with some justification, that this approach assumes that all academics can leverage enough power to exert sufficient influence to change the status quo. Certainly, well-established academics with a strong research and publication profile might have sufficient social capital and a high enough public visibility to be consulted on issues at a national or international scale. They might have sufficient capital to mobilize institutional and network resources that will open doors of their choosing and build relationships and goodwill. The majority of academics, however, have less clout and reputation to gain insider status; to do so requires work and its development will be gradual. The key is to start locally, with a realistic goal and a contribution that is meaningful and sustainable (in terms of time, resources, knowledge, ideas, and continued engagement over a span of years). Organizations at a local scale looking for informed help

and resources are keen to form partnerships and extend networks. Often seeing academics as useful allies, they appreciate the reputational effect of having university involvement in their initiatives as well as the potential access to shared resources through collaborative funding applications. Successful work at a local level tends to gain the attention of stakeholders at the next scale, with project work, “grey” publications, events, media work, and networks providing a means to cultivate relationships and develop sufficient capital to move up into a new position. Delivering on promises and building networks are vital to gain trust and new openings. Even if one stalls at a particular scale, one might hold a reasonable degree of sway at that level. I have followed this ladder route, starting by working with local authorities and civic organizations and then progressing upwards, eventually serving on national-level advisory boards (e.g., Data Forum, Dept of Taoiseach [Prime Minister]; Census Advisory Board; board of the Irish Research Council; National Consultative Panel on Open Data) and working with government departments and state agencies.

AIRO, a joint venture between Maynooth University and Dundalk Institute of Technology, was founded in 2005 as a means to produce harmonized cross-border datasets spanning the Republic of Ireland and Northern Ireland (see chapter 7 of Kitchin 2021). In the wake of the Good Friday Agreement and the peace process in Northern Ireland, cooperation between public sector bodies in the North and the Republic, including shared infrastructure and development plans, had increased enormously. However, there was a dearth of evidence that might underpin these endeavors and little activity to resolve the lacunae. Given animosity between political parties, the difficulty of organizing cooperation between public bodies in different jurisdictions, and entrenched statistical systems and geographies, the most viable solution was the creation of an independent third party to investigate possibilities for creating cross-border data infrastructure, including interactive data tools, for the planning, assessment, and tracking of cross-border developments. AIRO proposed to become that third party, with the Special EU Programs Body (SEUPB) providing the initial funding (€68,000). It was hoped that AIRO would be able to leverage the established applied GIS expertise and networks of its principal investigators to tackle the challenge.

While much of the project tasks were technical, a substantial aspect of the project work consisted of outreach and negotiation with stakeholders to convince them to work with the project and to embrace its vision and ambitions. Much of this advocacy work was highly political given the political sensibilities of key stakeholders concerning cross-border cooperation. It required the project team to act as mediators, to broker relationships, and

to lobby stakeholders and demonstrate the potential benefits and paths to achieving them. The team organized several events and bid for research contracts relating to data on both sides of the border and used these to further the cross-border data agenda. The project was also an early advocate of open data, insisting that the datasets and tools it had built for stakeholders were made openly available. Out of its small initial grant, AIRO built a sizable network of allies and a reputation for creating datasets and tools for evidence-informed policymaking, and it has subsequently undertaken work for all the local and regional authorities in the Republic and several in the North, as well as government departments and state agencies. It has placed researchers into state organizations to help develop their data practices and infrastructures along with their capabilities for data analytics and data-driven policy (see Gleeson et al., 2022). While remaining independent, it works closely with stakeholders to shape the Irish state's data regime. It can act in this way because it has worked hard to gain a trusted insider position. Nonetheless, much of what it advocates for remains unrealized given inertia and competing interests. In this sense, ongoing advocacy and change management work are always in play.

AIRO also provided part of the platform for splinter initiatives such as The Programmable City and Building City Dashboards projects. The Programmable City project, funded by the European Research Council, was principally concerned with undertaking a critical assessment of urban digital technologies and their use in city management and governance. This initiative included a focus on the production and use of open and proprietary data, as well as the deployment of data-driven systems. The project was to incorporate interviews with key stakeholders and ethnographies of smart city initiatives. Early in the project (2013), I was asked to join the steering committee of Dublinked, Dublin's open data repository, which was being jointly developed by the four local authorities and Maynooth University. In 2015, Dublinked was folded into the newly created Smart Dublin, a body for promoting smart city endeavors in the city, with the existing steering committee staying in place. In 2014, supplemental funding from Science Foundation Ireland initiated the production of a city dashboard for Dublin, to be developed in conjunction with Dublin City Council and leveraging foundational data auditing, collation, and visualization work by AIRO. The politics and praxes involved in the unfolding of this project are detailed in Kitchin et al. (2016). The project also aided Smart Dublin in some of its work by organizing and facilitating workshops to scope out the city's smart city vision and agenda. Later work prompted and helped develop its ethics approach to smart city technologies. In exchange for its contributions of

time, expertise, and resources, the project also gained access to key personnel for interviews. In fact, no exchange of funds ever took place between the university and the local authorities; the aid was mutual and *quid pro quo*. Through my involvement in the steering committee, the project's contribution to Smart Dublin, and the development of the city dashboard, I was able to occupy an insider position, directly contributing to open data and smart city policy and to the development of a data infrastructure. In 2016, Science Foundation Ireland funded the Building City Dashboards project that extended the dashboard partnership to include two local authorities in Cork and two state agencies (the Central Statistics Office and Ordnance Survey Ireland, the national mapping agency).

In these projects I, and others on my teams, have moved to varying degrees into insider positions. We have not simply been providing a research service, but have contributed to governance, policy formulation, agendas and visions, and the scope and workings of projects. Importantly, my work with the government has not been to do what they want or to say what they want to hear. I am supportive when support is merited, but I will also stick to principles, debating an issue and arguing a case, and being contrary when necessary. In contrast to my initial expectations, I have found that politicians and others in the public sector are fine with critical voices and robust discussion as long as this sort of engagement is backed with rhetoric and evidence, is constructive and fruitful, and their overall agenda is advanced. They are well used to internal debate and spats between colleagues and among cliques, and they have been subjected to public critique through the media. They are used to evolutions and mutations in the formulation of policy and its implementation, to the political context being relatively fluid, and to certain issues being politically hijacked. Nonetheless, critique and robust exchange can lead to some difficult situations if a partner is losing face or does not want to change the approach or if neither party wants to compromise. At the same time, the discernment of knowing which battles to fight and when to make a tactical concession or retreat is important. To continue to influence plans, decision-making, and actions, I want to stay inside the system rather than be frozen out. I am therefore prepared, as part of a longer game, to tolerate approaches, decisions, and policies that I think are suboptimal or regressive. This long-game orientation requires committed involvement over several years. My strategy is to put my opposition on the record through publication (meeting minutes, social media, blog posts, media interviews, academic articles) while not making a given matter a do-or-die issue. Such pragmatism will not suit all academics, who might feel that it would overly compromise their integrity, independence, and

impartiality. Ultimately, I feel I am more likely to gain a desired outcome by being inside the system. That is not to say that I will compromise on all issues, and if there was a decision or approach that I felt I could not tolerate being a party to I would actively oppose it and, if necessary, withdraw from the process.

Conclusion

The role of academics in society is very much a live debate within the academy and its disciplines, as well as within political, policy, and media circles. Some academics frame academic endeavor as the production of objective and value-free knowledge that others translate for instrumental ends. Others contend that academics should seek to change the world in proactive ways, and that all research and dissemination is inherently political. Some are concerned about the independence and impartiality of academic work, and the extent to which it is being enlisted in the agendas of other stakeholders. Nonetheless, academics are increasingly being asked to make a societal impact beyond their work being acknowledged and used by their peers. Academic researchers are being encouraged to engage the public through media channels and to work with civil society, industry, and government to address societal challenges. For post-positivist scholars, such collaborations pose certain concerns, since research rooted in critical social theory often challenges the ethos, rhetoric, and actions of government and industry and seeks to hold them to account. Working with these stakeholders necessarily means compromising on specific ideas and ideals. Yet insider positions are a powerful means of actively shaping the thinking, decision-making, and actions of stakeholders; and to exert influence from such a position is often more effective than critiquing from the outside. This chapter has sought to illustrate the concerns and praxis of such insider researchers in relation to three sustained engagements with data regimes in Ireland. Key to these projects and the effort to enact progressive change from within has been a keen sense of positionality and reflexivity (Rose 1997), both in relation to collaboration and in terms of publication stemming from the work. As an advocate-researcher, I have always sought to be open about my agenda, my situatedness within the research and policy fields, and any compromises that have been made. This transparency is important to help others frame and understand the knowledges being produced and the kinds of instrumental practice being enacted, and they help to negate, or at least make clear, concerns over independence and impartiality. While such concerns have

some legitimacy, it is my contention that using academic endeavor to change society for the better by enhancing data regimes through working with key stakeholders (who themselves are far from independent and impartial) as “critical friends” is vital.

References

- Allen, Chris. 2011. “Against Dialogue: Why Being Critical Means Taking Sides Rather Than Learning How to Play the ‘Policy Research’ Game.” *Dialogues in Human Geography* 1 (2): 223–227.
- Bastow, Simon, Patrick Dunleavy, and Jane Tinkler. 2014. *The Impact of the Social Sciences: How Academics and Their Research Make a Difference*. London: Sage.
- Burawoy, Micheal. 2005. “For Public Sociology.” *American Sociological Review* 70: 4–28.
- Dalton, Craig, and Jim Thatcher. 2014. “What Does a Critical Data Studies Look Like and Why Do We Care?” *Society & Space blog*, May 19, 2014. Accessed December 1, 2020. <http://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care>.
- Gleeson, Justin, Rob Kitchin, and Eoghan McCarthy. 2022. “Dashboards and Public Health: The Development, Impacts, and Lessons from the Irish Government COVID-19 Dashboards.” *American Journal of Public Health* 112 (6): 896–893.
- Fischer, Frank. 2005. “Beyond Empiricism: Policy Inquiry in Post-positivist Perspective.” *Policy Studies Journal* 26 (1): 129–146.
- Fuller, Duncan, and Rob Kitchin, eds. 2004a. *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?* Praxis E-Press. <http://www.praxis-epress.org/rtcp/contents.html>.
- Fuller, Duncan, and Rob Kitchin. 2004b. “Radical Theory/Critical Praxis.” In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 1–20. Vernon and Victoria: Praxis (e)Press.
- Kitchin, Rob. 2015. “Positivist Geography.” In *Approaches to Human Geography*, 2nd edition, edited by Stuart Aitken, and Gill Valentine, 23–34. London: Sage.
- Kitchin, Rob. 2021. *Data Lives: How Data Are Made and Shape Our Lives*. Bristol: Bristol University Press.
- Kitchin, Rob. 2022. *The Data Revolution: A Critical Approach to Big Data, Open Data, and Data Infrastructures*, 2nd edition. London: Sage.
- Kitchin, Rob, Denis Linehan, Cian O’Callaghan, and Philip Lawton. 2013. “Public Geographies and Social Media.” *Dialogues in Human Geography* 3 (1): 56–72.
- Kitchin, Rob, Sophia Maalsen, and Gavin McArdle. 2016. “The Praxis and Politics of Building Urban Dashboards.” *Geoforum* 77: 93–101.

- Leonelli, Sabina, Brian Rappert, and Gail Davies. 2017. "Data Shadows: Knowledge, Openness, and Absence." *Science, Technology & Human Values* 42 (2): 191–202.
- Markham, Annette N. 2017. "Troubling the Concept of Data in Digital Qualitative Research." In *Handbook of Qualitative Data Collection*, edited by Uwe Flick, 511–523. London: Sage.
- Mitchell, Don. 2004. "Radical Scholarship: A Polemic on Making a Difference Outside the Academy." In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 21–30. Vernon and Victoria: Praxis (e)Press.
- Mouton, Johann, and Hendrik Christoffel Marais. 1988. *Basic Concepts in the Methodology of the Social Sciences*. Pretoria: HSRC Publishers.
- Rose, Gillian. 1997. "Situating Knowledges: Positionality, Reflexivity and Other Tactics." *Progress in Human Geography* 21, 305–320.
- Wilton, Robert. 2004. "Keeping Your Distance: Balancing Political Engagement and Scientific Autonomy with a Psychiatric Consumer/Survivor Group." In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 116–131. Vernon and Victoria: Praxis (e)Press.

About the Author

Rob Kitchin is a professor in Maynooth University Social Sciences Institute, for which he was director 2002–2013 and 2021–22. He is currently an ERC Advanced Investigator for the project 'Data Stories'. He is the (co)author/editor of 35 academic books and (co)author of over 200 articles and book chapters. He is a recipient of the RIA Gold Medal for the Social Sciences.

3. **Confronting Politicized Research: The Case for Reflexive Neutrality**

René König, Payal Arora, and Usha Raman

Abstract

We have come a long way from a model that idealizes politically neutral research to one that expects researchers to take a stand and to actively change society. Our paper retraces this development and highlights the criticisms that led to this change in orientation. Concluding that it might be too early to entirely abandon the idea of neutrality, we outline initial ideas of a revised model that considers certain valid criticisms while maintaining the goal of neutrality.

Keywords: Knowledge production; Epistemology; Objectivity; Post-truth era

Introduction

Increasingly researchers are expected to ‘make a difference,’ an appeal that a few decades ago would have been deeply controversial in its contention that scholars should actively shape society. Then the prevalent mindset of positivism expected that researchers would be neutral observers, merely dedicated to truth and facts, instead of wading into the muddy waters of activism and politics. Today, this formerly accepted approach is widely regarded as naïve. Constructivism has made us question the very existence of an objective reality; science and technology studies have taught us about the social embeddedness of research practices, and theoretical and methodological discussions have highlighted the limitations of academic knowledge production. Yet the imposed binary between the passive model of scholars as neutral observers and the active model of scholars as agents of social

change should make us pause. Before we dive headfirst into the endeavor of ‘making a difference,’ it may be worth our while to reflect on how we got here. In this essay, we take a look at what caused this significant turnaround of scholarly principles. We then critically assess the implications of the new expectations directed towards researchers and the accompanying claims about their role in society. Finally, we conclude with a provocation: a call to revive the ideal of academic neutrality – in a modified and reflective way.

The Neutral Researcher – From Idealization to Disillusionment

The ideal of objective research, not motivated by particular political intentions and biases but solely dedicated to truth-seeking, is strikingly ambivalent: on the one hand, the aforementioned decades of critical philosophical discussions, methodological reflections, and empirical insights on how research is performed, make proponents of this ideal appear, at best, uninformed and, at worst, willfully manipulative. On the other hand, science certainly has achieved an elevated status and a special authority in the hierarchy of knowledge types. As Böhme and Stehr argued, “contemporary society may be described as ‘knowledge society’ based on the penetration of all its spheres of life by scientific knowledge” (1986, 10:8).

To some degree, the resistance and skepticism towards scientific knowledge does not *contradict* but rather *result from* the very success of this mode of knowledge production. From the devastation wrought by the world wars, increasing ecological destruction, and, finally, the potential total annihilation of humanity in the nuclear age, the twentieth century was a testament to the deep but also troublesome impact of supposedly ‘objective’ scientific and technological development. The growing awareness of modern societies’ self-induced risks led scholars to diagnose a “reflexive modernity” (Beck 1992; Giddens 1990). Science and technology were no longer regarded as a solution but increasingly viewed as a problem.

Moreover, persistent inequality and poverty meant that the promised benefits of scientific progress were unevenly distributed. This disparity became not just a question of justice and fairness but also of scientific validity. How can universal and objective knowledge possibly be produced if the access to the necessary tools and skills is limited to a small but dominant elite? Accordingly, feminist scholar Donna Haraway described the claim of scientific objectivity and universality as the “god trick of seeing everything from nowhere” (1988, 581). Instead, she emphasized the situatedness of knowledge and its relation to power structures (1988, 585). Noting that

so-called objectivity “may predicate colonial practices and establish a hegemonic ideology of difference” decolonial studies researchers concluded “[a]ll research is political” (Sandoval et al. 2016, 27).

The existential threats posed by scientific knowledge and the uneven means to produce and benefit from it are particularly challenging for democratic societies, whose egalitarian aspirations have long conflicted with the elitist model of scientific expertise (Fischer 2009). Accordingly, there have been numerous appeals for a deliberate blurring of boundaries beyond the confining categories of scientific disciplines (Gibbons et al. 1994) and towards a ‘democratization’ of science and technology (Hennen 2012; Guston 2004; Nowotny 2003). The rise of the internet, particularly the platform-driven so-called Web 2.0 (O’Reilly 2005), in combination with the wide availability of smartphones, have fueled hopes of a more inclusive “wisdom of crowds” in which a wider public participates in knowledge production (Surowiecki 2004; Rheingold 2003; Shirky 2008).

Of course, not everyone welcomed this transformation. Fears of a counterproductive ICT-related “information overload” had already been voiced as far back as in 1970 (Toffler 1970), and commentators have criticized the novel influence of amateurs (Keen 2007). Moreover, it has become obvious that the emerging “platform society” (Van Dijck, Poell, and Waal 2018) is itself entangled in power structures and political struggles and is largely driven by capitalist ideology (Zuboff 2019; Fuchs 2011; Mager 2012).

Thus, seemingly ‘democratizing’ developments may cement existing hierarchies rather than empowering the marginalized. For example, the open science movement started with the goal to “democratize all aspects of the scientific process, from conceptualization and design, to data gathering, to publication and peer review, to dissemination of research findings” (Dutta et al. 2021, 2). However, as Dutta et al. point out, power imbalances are in practice actually deepened:

The steps to achieve openness, including pre-registration of studies, publishing materials, data and code, and conducting replications, add exponential burdens to the already complicated submission procedures, U.S.-based English language hegemony, and inaccessibility behind paywalls that shape the hierarchy of publishing [...] (ibid., 5).

Additionally, opening research practices up may disproportionately harm marginalized and vulnerable communities (Fox et al. 2021, 767). This potential outcome is particularly concerning, as researching such groups is inherently burdensome for participants as well as researchers, bringing into

relief the fact that even as science cannot be objective and disinterested, neither can the researcher. Rawshon Akhter, a colleague from our research project *Feminist Approaches to Labor Collectives (FemLab.Co)*, recalls a dramatic incident during her fieldwork:

[...] a divorced mother of three young daughters became faint when she recounted her experience of domestic abuse meted out to her for failing to give birth to a male child. I had to call for help as she was close to collapsing. (Akhter 2020)

Such vivid examples underline the artificiality and unrealistic nature of the ideal of the objective and neutral researcher. Observation never occurs in isolation. Empirical studies may exert a profound impact on the field and its people, and this has implications for the researcher. Based on her experience, Akhter reflects: “Researchers are not psychologists nor social workers. However, they may find themselves playing these roles unconsciously as they try to be of help to their respondents” (ibid.).

These concerns raise the questions: What, exactly, is the researcher’s role if it is not that of a neutral observer? If research cannot be neutral, what stance is the researcher supposed to take?

Research in the Post-Truth Era

Objectivity and neutrality are by nature ideals vulnerable to criticism, as they require preconditions that are hard – if not impossible – to meet. However, finding a convincing alternative model is no less challenging. Debates on how normative research can or should be have gone on for decades, especially in the social sciences, where it is particularly difficult to draw a line separating researchers and their object of study. In German sociology, Niklas Luhmann and Jürgen Habermas engaged in a prominent disagreement about the discipline’s purpose (Habermas and Luhmann 1990). Habermas made the case for a rather ‘interventionist’ role, a perspective rooted in the critical theory of the Frankfurt School. Luhmann, for his part, was convinced that this normative approach would be limiting, as it would bind the discipline to certain assumptions instead of aiming to formulate a general theory and understanding of society.

Although this theoretical debate was mostly concerned with the narrow disciplinary circle of sociology, wider audiences, too, have debated the actual or alleged normative biases among (social) scientists. In 2018,

Helen Pluckrose, James A. Lindsay, and Peter Boghossian effectively hoaxed numerous academic journals when they submitted a series of fake papers, including one that was actually a chapter of Adolf Hitler's *Mein Kampf*, though it had been rewritten from a supposedly feminist perspective (and its source text left undisclosed). Their motivation was to reveal the biases of some parts of academia they labeled "Grievance Studies". They stated:

Scholarship based less upon finding truth and more upon attending to social grievances has become firmly established, if not fully dominant, within these fields, and their scholars increasingly bully students, administrators, and other departments into adhering to their worldview. This worldview is not scientific, and it is not rigorous. (Pluckrose, Lindsay, and Boghossian 2018)

Pluckrose et al. saw their point confirmed when some of the papers got through the peer review process, including the rewritten chapter from *Mein Kampf*. The story of the hoax was picked up by several larger outlets, including *The Atlantic* (Mouk 2018) and *The Wall Street Journal* (Melchior 2018). The incident mirrors the so-called 'Sokal Affair' or 'Sokal Hoax' of 1996, when physicist Alan Sokal successfully published a bogus article to the journal *Social Text* to demonstrate the journal's ideological bias and lack of rigor. At the core of his criticism was what he saw as the denial of objective truth among postmodern scholars and the inappropriate usage of scientific terminology.

Lee McIntyre has described the denial of objective truth as the "first thesis of postmodernism". A second thesis, according to him, is that "any profession of truth is nothing more than a reflection of the political ideology of the person who is making it" (McIntyre 2018, 126). Referring to Michel Foucault, he concludes his sharp criticism of postmodern thought by drawing out the implications of its underlying logic:

Since there is no such thing as 'truth,' anyone who claims to 'know' something is really just trying to oppress us, not educate us. Having power allows us to control what is true, not the other way around. If there are many perspectives, then insisting that we accept any particular one is a form of fascism. (ibid.)

As McIntyre himself notes, this account is a rather condensed and perhaps an unfair and distorted interpretation of postmodern thought. We do

acknowledge the rich and enlightening research inspired by this sort of postmodernism; in fact, our own work is built on it. The FemLab.Co project is informed and driven by the insights offered via the lens of critical theory, e.g., feminist and decolonial studies. Our project team has embraced feminist design as an approach to platform workers' problems (Bansal 2021), applied feminist methodology in creating an inclusive workshop (de Souza and Gupta 2021), drawn from a critical studies perspective to reveal the gaps of methods such as those employed in stakeholder analysis (Mehta 2021), to name just a few examples.

Therefore, we are fully aware that the cited criticisms – let alone our concentrated version presented here – do not do justice to such a complex and multifaceted research tradition. Moreover, these critiques are themselves not without flaws. For instance, the “grievance studies” hoax was, ironically, accused of itself adhering to what it intended to reveal: an unscientific, ideology-driven approach (Afinogenov 2018).

However, it would be careless to simply dismiss the raised concerns – especially if we value the ideas being criticized. When the political value and the potential societal impact of academia become decisive factors in research, there is an inherent risk that the traditional focus on truth-seeking may suffer. Quality standards may be lowered due to political motives or political motives may create barriers to the execution of high-quality research. In short, ideology may be prioritized over truth. Not only does this diminish the original function of scientific inquiry, but such efforts will often backfire, as their underlying motivation detracts substantially from the power such research could have had. Thus even from a political standpoint, where the primary aim is to aid a particular cause or a group, the adoption of this strategy does not promise desirable results. It does, however, leave us with a conundrum to address.

If all research is inherently political, we need to ask: what political agenda is – or should be – pushed here? Who is defining such an agenda and how? So far, academia has not been very well equipped for political decision-making. Instead, political goals often remain implicit or vague. To some extent, the schools emerging from critical studies share certain progressive ideas aimed at emancipating marginalized groups and questioning oppressive structures – an approach rooted in the 18th-century Enlightenment. Still, it is by no means clear *which* values are deemed suitable. Even a seemingly shared ideology – for example, feminism – often possesses many diverging schools of thought (Delmar 2018). Thus, defining the values that are supposed to guide a research project and translating them into methods and practice are far from trivial tasks.

In an often cited introductory book on action research, Jean McNiff lays this responsibility largely at the foot of researchers, since they are asked to “clarify your values from the start” (McNiff 2016, 20). In the validation process, however, this perspective is expanded, as it is supposed to involve “social validation, to do with testing the validity of the claims against the critical feedback of others” and “public legitimation.” While social validation should help to assess “truthfulness,” public legitimation is aimed at “establishing the acceptability of the claim” (McNiff 2016, 52).

While this approach appeals to the idea of the ‘democratization’ of science, it raises numerous additional questions. How independent can or should research be? On the one hand, independence is emphasized by letting researchers define their values. On the other hand, these researchers are then asked to let the public decide whether their findings should be accepted.

One underlying assumption behind the idea of welcoming the public into academia appears to be the assurance that the public will show itself to be on the ‘right’ side – or even on some ‘better’ side that researchers have neglected. Here we find the postmodern version of the “god trick”: instead of researchers being glorified as objective observers, the public is imbued with an elevated function and is idealized. But it is not clear why the public should be graced with a superior perspective that would justify such an idealization. After all, each member of the public is entangled in individual power structures, biases, and distorted perceptions of their own. There is no reason to assume that these perspectives will be compatible with a given research project’s values. Would that render the results ‘illegitimate’? If not, why involve the public at all?

For us, this is not just a hypothetical discussion. In the course of engaging with fieldwork for our project FemLab.Co, we were often confronted with regressive and misogynist views. In fact, such views are what motivated us to take a specifically feminist perspective on the Future of Work in the Global South. Women in countries such as India and Bangladesh are exposed to discrimination and injustice on a daily basis, and thus the public is a central part of the problem we are trying to tackle. Our researchers and participants often need *protection* from the public, not its active involvement.

In 2020, only 8.4 percent of the world population lived in full democracies (Economist Intelligence Unit 2021). If the democratization of political systems is so unsuccessful, how realistic is it to hope for a democratization of science in accordance with progressive values? To make matters worse, even in full democracies it seems unlikely that the high hopes placed on citizen participation will be met. In 2017, Kellyanne Conway, then Senior Counselor to US president Donald Trump, famously coined the term “alternative facts”.

Her phrase and what it implied caused a fair amount of bewilderment among those who still believed in objectivity, and they naturally turned towards academia as an ally to ward off the looming ‘post-truth’ era. In his book on the phenomenon, McIntyre points the finger at postmodernism and states that it is “embarrassing to admit that one of the saddest roots of the post-truth phenomenon seems to have come directly out of colleges and universities” (2018, 123). We cannot know if Conway and her peers have read postmodern theory, of course. But presuming that there are multiple versions of the ‘truth’ is certainly compatible with postmodern perspectives. This incident made it painfully obvious how the acceptance of truth in a plural form may serve the powerful and promote regressive ideas. This is not a new insight. McIntyre explains in his book (2018) how tobacco companies have tried to establish scientific ‘facts’ on the products’ health hazards, and how conservatives have created counternarratives to the overwhelming scientific consensus on anthropogenic global warming. More recently, ‘anti-vaxxing’ movements have challenged established scientists and their point of view on how the COVID-19 pandemic should be understood and fought. These challenges exert severe impacts not only on public health but also on the scientists themselves, who increasingly face negative consequences after public appearances, including death threats (Nogrady 2021). The direct link between academics and the public through social media has made such experiences almost unavoidable in certain fields: researchers now have to find their own coping strategies to deal with this problem (Gewin 2018).

What can we learn from all this? Researchers cannot be neutral. They do not operate in a social vacuum. But we should not then assume that the politicization of research is the answer here. At the very least, we need to acknowledge that politicization is not a solution to the underlying epistemological and societal problems in the context of scientific expertise. Perhaps letting go of the traditional ideal of academic neutrality was premature, even careless.

Conclusion: Towards a Reflexive Neutrality

While Nazis in Germany and elsewhere attempted to justify their murderous policies through racist pseudoscience, the sociologist Robert K. Merton reflected on what the foundations of the underlying ethos of science should be. Among the “institutional imperatives” (1973, 270 ff.) invoked by Merton is a stance that appears to be the opposite of our contemporary, deliberately political approaches to research: *disinterestedness* (Merton 1973, 275). Merton

saw the goal of science in the “extension of certified knowledge” (Merton 1973, 270), something he believed could be achieved via an institutional and functional detachment from society and the partisan goals of its political actors. In contrast to the later idealization of a ‘democratization’ of science, this perspective regards scientific independence and autonomy as beneficial for both sides: a disinterested orientation helps to resist both the politicization of knowledge production and its instrumentalization *vis-à-vis* a public that may not be equipped to judge it.

The lessons learned in the phase of reflexive modernity preclude idealization of the ‘ivory tower’ – science situated outside and somewhat above society and its concerns. In any case, such an ideal would be to no avail. Science and the knowledge society are closely interconnected and the advent and now the ubiquity of the internet has made the ‘ivory tower’ irreversibly more accessible (Nentwich and König 2012, 11:151 ff.). We should not, however, be misled to idealize the opposite stance: an intentional politicization of science and academia.

Perhaps we are entering the era of *reflexive postmodernity*. During the era of reflexive modernity, society dealt with the risks from modernity (see above), but now we face the self-induced risks of postmodern society. One strategy to cope with our predicament might be to re-establish academia as a depoliticized safe space for free thought, adhering to the ideal of uncovering the truth in its singular form. As McIntyre has put it: “There is no such thing as liberal science or conservative science. When we are asking an empirical question, what should count most is the evidence” (2018, 163).

Accordingly, we argue that it is time to revisit the concept of academic neutrality, while also considering previous criticisms. First and foremost, we must acknowledge that *neutrality is a process*. It is a never-fully-achievable ideal. As such, it requires constant reflection on one’s values, biases, and methods (see also Markham 2006). The ideal’s inherent unattainability also implies that no researcher can claim absolute neutrality, as the “god trick” allowed them to [claim]. Instead, our ideal of *reflexive neutrality* does not deny the social embeddedness of academia, shot through with its political and cultural particularities.

All research is political – but not all researchers are politicians, nor need they be. Indeed, researchers can choose how they position themselves in politically contested fields. As Roger Pielke Jr. (2007) suggests, they may act the role of an “honest broker”, making diverging perspectives visible instead of negating them under the cover of an alleged objective truth. This does not mean that researchers need to present all perspectives evenly and without judgement. In fact, in many cases, to do so would result in an endless and

ultimately meaningless list of random claims. A core task in research is always to distinguish between valid and invalid claims.

Reflexive neutrality is not some utopian ideal. One variation of reflexive neutrality is ongoing and has been fairly successful: Wikipedia. The online encyclopedia has instituted a Neutral Point of View (NPOV) policy for its articles, which is constantly reflected upon through a collaborative process. Various critics have argued that the platform is not actually neutral (Lovink and Tkacz 2011). Indeed, a bias towards scientism can be identified, for example, in the way the platform handles unorthodox claims such as conspiracy theories (König 2012). But this would only be a problem if the goal was an entirely agnostic and radical neutrality. Reflexive neutrality, however, acknowledges its biases – here, scientism – and even deliberately nurtures them, as long as they are justifiable. At the same time, there needs to be an openness to legitimate criticism. For example, decolonization scholars have argued that Wikipedia’s ‘No Original Research’ policy “creates barriers for those consistently underrepresented in writing or publishing: historically marginalized communities and knowledge systems relying on oral traditions” (Acey et al. 2021, 3). However, such biases are not an argument against neutrality but rather indications of its inadequate application.

Admittedly, our appeal for reflexive neutrality provides more questions than answers. We do hope, though, that we have raised questions that make the reader pause on the quest to ‘make a difference’ – because we do share the urge to affect societal change. Certainly, many questions remain that deserve further thought: How can reflexive neutrality be achieved on a practical level? What are the necessary organizational structures and methodologies? More fundamentally, the role of science in society needs to be discussed beyond academia and expert circles. Or, to change the perspective, the problem of politicized research is also one of rationalized politics. In opposition to the popular notion of ‘post-truth’, Alexander Bogner (2021) puts the power of knowledge at the center of concern. From this point of view, the dominance of science has led to a depoliticization of political discourse, often framing it in terms of questions of rationality and truth instead of values and justice.

While the idea of reflexive neutrality is difficult to outline in satisfying detail, the model of a deliberately non-neutral researcher presents us with even more questions to answer and more problems to solve. Rather than imagining some unachievable clarity, we should acknowledge the complexity and muddiness that unavoidably evolves when science and politics mingle. Taking this reality seriously and learning from past mistakes, we can only conclude that we are compelled to engage in ongoing reflection.

Acknowledgements:

This chapter is part of the FemLab.Co project funded by the International Development Research Center (IDRC), Project Number-Component Number: 109331-001.

References

- Acey, Camille E, Siko Bouterse, Sucheta Ghoshal, Amanda Menking, Anasuya Sengupta, and Adele G Vrana. 2021. "Decolonizing the Internet by Decolonizing Ourselves: Challenging Epistemic Injustice through Feminist Practice." *Global Perspectives* 2 (1): np.
- Afinogenov, Greg. 2018. "Orthodoxxed! On 'Sokal Squared.'" *n+1* (blog). October 4, 2018. <https://www.nplusonemag.com/online-only/online-only/orthodoxxed/>.
- Akhter, Rawshon. 2020. "Research That Hurts – On the Practice of Care in Fieldwork." *FemLab.Co* (blog). June 26, 2020. <https://femlab.co/2020/06/26/research-that-hurts-on-the-practice-of-care-in-fieldwork/>.
- Bansal, Pallavi. 2021. "Is Feminist Design a Solution to Platform Workers' Problems?" *FemLab.Co* (blog). February 20, 2021. <https://femlab.co/2021/02/20/is-feminist-design-a-solution-to-platform-workers-problems/>.
- Beck, Ulrich. 1992. *The Risk Society: Towards a New Modernity*. London: Sage.
- Böhme, Gernot, and Nico Stehr. 1986. *The Knowledge Society: The Growing Impact of Scientific Knowledge on Social Relations*. Vol. 10. Dordrecht: Springer Science & Business Media.
- Bogner, Alexander. 2021. *Die Epistemisierung des Politischen. Wie die Macht des Wissens die Demokratie gefährdet*. Stuttgart: Reclam.
- Delmar, Rosalind. [2001] 2018. "What Is Feminism?" In *Theorizing Feminism: Parallel Trends in the Humanities and Social Sciences*, edited by Anne C. Herrmann and Abigail J. Stewart, 5–28. London: Routledge.
- Dijck, José van, Thomas Poell, and Martijn De Waal. 2018. *The Platform Society: Public Values in a Connective World*. New York, NY: Oxford University Press.
- Dutta, Mohan, Srividya Ramasubramanian, Mereana Barrett, Christine Elers, Devina Sarwatay, Preeti Raghunath, Satveer Kaur, Debalina Dutta, Pooja Jayan, and Mahbubur Rahman. 2021. "Decolonizing Open Science: Southern Interventions." *Journal of Communication* 71 (5): 803–826.
- Economist Intelligence Unit. 2021. "Democracy Index 2020: In Sickness and in Health?" <https://www.eiu.com/n/campaigns/democracy-index-2020/>.
- Fischer, Frank. 2009. *Democracy and Expertise: Reorienting Policy Inquiry*. New York, NY: Oxford University Press.

- Fox, Jesse, Katy E. Pearce, Adrienne L. Massanari, Julius Matthew Riles, Łukasz Szulc, Yerina S. Ranjit, Filippo Trevisan, Cheryll Ruth R. Soriano, Jessica Vitak, and Payal Arora. 2021. "Open Science, Closed Doors? Countering Marginalization through an Agenda for Ethical, Inclusive Research in Communication." *Journal of Communication* 71 (5): 764–784.
- Fuchs, Christian. 2011. "Web 2.0, Prosumption, and Surveillance." *Surveillance & Society* 8 (3): 288–309.
- Gewin, Virginia. 2018. "Real-Life Stories of Online Harassment –and How Scientists Got through It." *Nature* 562 (7727): 449–450. <https://doi.org/10.1038/d41586-018-07046-0>.
- Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. 1994. *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage.
- Giddens, Anthony. 1990. *Consequences of Modernity*. Cambridge: Polity Press.
- Guston, David H. 2004. "Forget Politicizing Science. Let's Democratize Science!" *Issues in Science and Technology* 21 (1): 25–28.
- Habermas, Jürgen, and Niklas Luhmann. 1971. *Theorie der Gesellschaft oder Sozialtechnologie*. Frankfurt am Main: Suhrkamp.
- Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–599.
- Hennen, Leonhard. 2012. "Why Do We Still Need Participatory Technology Assessment?" *Poiesis & Praxis* 9: 27–41. <https://doi.org/10.1007/s10202-012-0122-5>.
- Keen, Andrew. 2007. *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today's User-Generated Media Are Destroying Our Economy, Our Culture, and Our Values*. New York, NY: Doubleday.
- König, René. 2012. "Wikipedia. Between Lay Participation and Elite Knowledge Representation." *Information, Communication & Society*, October. <https://doi.org/10.1080/1369118X.2012.734319>.
- Lovink, Geert, and Nathaniel Tkacz, eds. 2011. *Critical Point of View: A Wikipedia Reader*. Amsterdam: Institute of Network Cultures. <http://networkcultures.org/wpmpu/portal/publications/inc-readers/critical-point-of-view-a-wikipedia-reader/>.
- Mager, Astrid. 2012. "Algorithmic Ideology. How Capitalist Society Shapes Search Engines." *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2012.676056>
- Markham, Annette N. 2006. "Ethics as Method." *Journal of Information Ethics*, 15(2): 37–54. DOI:10.3172/JIE.15.2.37.
- McIntyre, Lee. 2018. *Post-Truth*. Cambridge, MA: MIT Press.
- McNiff, Jean. 2016. *You and Your Action Research Project*. London: Routledge.
- Mehta, Chinar. 2021. "What Does the Abandoned Public Toilet Tell Us?" *FemLab. Co* (blog). 2021. <https://femlab.co/2021/05/21/what-does-the-abandoned-public-toilet-tell-us/>.

- Melchior, Jillian Kay. 2018. "Opinion | Fake News Comes to Academia." *Wall Street Journal*, October 5, 2018, sec. Opinion. <https://www.wsj.com/articles/fake-news-comes-to-academia-1538520950>.
- Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago Press.
- Mounk, Yascha. 2018. "What an Audacious Hoax Reveals About Academia." *The Atlantic*. October 5, 2018. <https://www.theatlantic.com/ideas/archive/2018/10/new-sokal-hoax/572212/>.
- Nentwich, Michael, and René König. 2012. *Cyberscience 2.0. Research in the Age of Digital Social Networks*. Vol. 11. Interactiva. Frankfurt: Campus.
- Nogrady, Bianca. 2021. "'I Hope You Die': How the COVID Pandemic Unleashed Attacks on Scientists." *Nature* 598 (7880): 250–253. <https://doi.org/10.1038/d41586-021-02741-x>.
- Nowotny, Helga. 2003. "Democratising Expertise and Socially Robust Knowledge." *Science and Public Policy* 30 (3): 151–156.
- O'Reilly, Tim. 2005. "What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software." September 30, 2005. <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>.
- Pielke Jr, Roger A. 2007. *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge: Cambridge University Press.
- Pluckrose, Helen, James A. Lindsay, and Peter Boghossian. 2018. "Academic Grievance Studies and the Corruption of Scholarship." *Areo*. October 2, 2018. <https://areomagazine.com/2018/10/02/academic-grievance-studies-and-the-corruption-of-scholarship/>.
- Rheingold, Howard. 2003. *Smart Mobs: The Next Social Revolution*. Revised edition. Cambridge, MA: Basic.
- Sandoval, Cueponcaxochitl D. Moreno, Rosalva Mojica Lagunas, Lydia T. Montelongo, and Marisol Juárez Díaz. 2016. "Ancestral Knowledge Systems: A Conceptual Framework for Decolonizing Research in Social Science." *AlterNative: An International Journal of Indigenous Peoples* 12 (1): 18–31.
- Shirky, Clay. 2008. *Here Comes Everybody: The Power of Organizing Without Organizations*. London: Penguin.
- Souza, Siddharth de, and Siddhi Gupta. 2021. "Collective Brainstorming to Imagine a Feminist Labour Collective." *FemLab.Co* (blog), 2021. <https://femlab.co/2021/10/04/collective-brainstorming-to-imagine-a-feminist-labour-collective/>.
- Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York, NY: Anchor Doubleday.
- Toffler, Alvin. 1970. *Future Shock*. New York, NY: Random House.
- Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. London: Profile Books.

About the Authors

René König is a sociologist and researcher at Rhine-Waal University of Applied Sciences and a PhD student at the University of Stuttgart. He is interested in the intersection of science and society and transformation processes related to the internet. He has worked for over ten years in the interdisciplinary field of technology assessment.

Payal Arora is a Professor of Inclusive AI Cultures at Utrecht University, and Co-Founder of FemLab, a future of work initiative. She is a digital anthropologist and author of award-winning books including 'The Next Billion Users' with Harvard Press. She has held fellow positions at GE, ZEMKI, ITRio, NYU and Rockefeller Bellagio.

Usha Raman is a Professor in the University of Hyderabad's Department of Communication, where her teaching and research focus on digital culture, feminist media studies, civic engagement, and media pedagogy. She has published over 30 research papers and book chapters and is the author of *Writing for Media* (Oxford University Press, 2009).

4. Inside Datafication: Entrepreneurial Research for Investigating Emerging Data Practices

Mirko Tobias Schäfer

Abstract

At Utrecht University, the Data School embarked on a journey to investigate AI and data practices up close, from within the organizations that implement them. This empirical work thrives on a process involving entrepreneurial activities – practices of collaborative and co-created research enabling researchers to deeply immerse themselves in the societal sectors under investigation. Occupying this privileged position, they do not only gather insights: they can also intervene and immediately apply their findings. This chapter presents how entrepreneurial activities are used for collecting evidence, creating learning opportunities, and possibilities for intervention and social impact. Referring to examples of entrepreneurial research at the Data School, the chapter discusses challenges and opportunities of this approach, and presents practices for mitigating risks.

Keywords: Participatory observation; Action research; Methods; Social impact

Introduction: Land and Expand

In response to the increasing need to use data from social media platforms for analyzing social interaction and political debates online, I co-founded the Data School (formerly Utrecht Data School) with Thomas Boeschoten – then a master’s student in Utrecht University’s media studies program – in 2013. As

there was no funding for building a course on digital methods – which was seen as too labor-intensive – we turned to external partners for financing. They would provide a team of students with a case or an issue related to datafication or their current data practices or those in development. They not only financed the additional contact hours needed for teaching but granted our students and researchers access to their organizations and data.

While their funding was helpful in the development of our program, we quickly noticed that there was a more appealing incentive for this kind of cooperation (Schäfer and Van Schie 2019). Students and researchers became embedded researchers who developed close cooperative relationships with members of the external partners' staff. Working with external partners allowed us to enter their organizations and the respective societal sectors, and thus provided a means to investigate the process of datafication and algorithmization from 'within.' By 'datafication' we refer to the process of translating everyday activities and social interactions into tabulated information (Mayer-Schönberger and Cukier 2013; Van Dijck 2014). By 'algorithmization' we refer to the use of models, machine learning, or what is generally called artificial intelligence (AI). But both 'datafication' and 'algorithmization' also carry ideological connotations of objectivity and accuracy claims, which exert strong effects on the strategy and management of organizations and profoundly shape the technological imaginaries related to Big Data and AI (e.g., Kennedy, People, and Van Dijck 2014; Richterich 2018; Männiste 2022).

To learn more about how AI and data practices affect democracy and citizenship, we looked for ways to conduct research within public management and media industries – areas where citizenship and democracy manifest themselves and where we could observe particular transformations. We developed this practice further so that it evolved into what we labeled *entrepreneurial research*, a distinctive method of empirical and collaborative research. This approach led to invaluable insights and to mutual knowledge transfer (e.g., Schäfer 2018; Schäfer, Van Es, and Muis 2023).

What is Entrepreneurial Research?

The ecstatic side of our fieldwork was experienced when the single-minded pursuit of data within a clearly defined research agenda was momentarily set aside, and the opportunity to enter deeply into the world of our hosts was embraced. To our surprise, this led to insights and knowledge that redefined the relationships with our hosts, deepened our ability

to interact with them in more meaningful ways, and opened the door to epistemological and ontological issues that begged to be addressed. (Goulet and Miller 2007, 1)

Originally, “entrepreneurial research” indicated research into entrepreneurship (e.g., Ucbasaran, Westhead, and Wright 2001; Grant and Perren 2002; Perren and Ram 2004; Landström and Lohrke 2010). However, we consider entrepreneurship to be not the object of our research but a central aspect of our research method. Furthermore, our concept of entrepreneurial research is not to be confused with “academic entrepreneurship” (e.g., Etzkowitz 2003). Academic entrepreneurship mostly describes individual researchers or research groups (often from the so-called STEM subjects) as well as the emergence of campus-embedded startups or university spinoffs (Shane 2004; Wright 2007; Pattnaik and Pandey 2014). It often provides accounts of the commercial application of research findings, the development of viable services and products on the basis of academic research, and the establishment of companies initiated by the commercial exploitation of academic research outcomes in a business model (Shane 2004); however, non-commercial initiatives, such as the example of social and humanistic entrepreneurship, can also fall in this category (Etzkowitz 2014).

The concept of academic entrepreneurship is also applicable to the far-reaching cooperation of university divisions with commercial parties. Universities with strong engineering and technology departments maintain close contacts with industry partners and take part in joint-venture research projects as well as other forms of cooperation and direct collaboration (Lee 1996, 2000). Pharmaceutical research, the health sciences, and agricultural research in particular have long traditions of cooperating with related branches of industry and of collaborating in research projects and in the development of products and services (see for example Powell, Koput, and Smith-Doerr 1996; Stuart 2000; Dooley and Kirk 2007).

One success story of academic entrepreneurship in the humanities is detailed by Pilegaard, Moroz, and Neergaard (2010), who describe an academic entrepreneur producing medical dictionaries. These efforts eventually led to the creation of a knowledge communication lab at Aarhus University that utilized research results commercially and attracted external funding (Pilegaard, Moroz, and Neergaard 2010, 52–53). Pilegaard et al. recognize the possibilities of entrepreneurship for educational and scientific innovation. Their main efforts, however, are directed towards the creation of better methodologies for researching entrepreneurship, rather than investigation of the benefits of entrepreneurial methods for academic research itself. Nevertheless, there

are calls for the humanities to develop applied research and interdisciplinary collaboration and to engage with societal partners (e.g., Brom 2019).

In his programmatic post “Towards the 4th Generation University,” Maarten Steinbuch (2016) describes the kind of transdisciplinary research that we can identify with: joint research collaborations of universities and external partners, carried out by academics and practitioners alike, that aim at a more dynamic and mutual knowledge exchange between the related organizations and participants, focused on basic research as well as the application of research findings. A similar perspective has been promoted by the computer scientist Ben Shneiderman, who claims that collaborations across disciplines and with partners outside the academe, as well as the combination of applied and basic research, will respond to the urgent research agenda better than traditional academic research has (Shneiderman 2016). While applied research was for too long a byproduct of academic research, their positions will now be reversed, and applied research will lead to basic academic research (Steinbuch 2016; Shneiderman 2026). Unsurprisingly, these voices come from the engineering and science departments, a more frequent site of interdisciplinary collaboration and cooperation with external partners than one finds in the humanities.

One reason for the scarcer occurrence of such cooperation in the humanities has been suggested by Andrew J. Nelson (2005) and his research on the entrepreneurial efforts of Stanford University’s musicology department. Nelson proposes that the logics of academia, which often include the sharing of knowledge in order to receive recognition, are difficult to merge with an entrepreneurial and commercial logic, which favors keeping insights to oneself so that one can financially exploit them. We suspect this argument is flawed. Our practice of entrepreneurial research depends on teamwork, collaboration across academic disciplines, and cooperation with external partners. The absence of incentives for doing this kind of labor-intensive and time-consuming, empirical, and transdisciplinary work, we would argue, is a far more plausible explanation.

With entrepreneurial research, we refer to a creative process of opportunity-seeking and risk-taking, which includes the use of commercial activities, among other means, to carry out research. Our research activities make use of entrepreneurial skills to enter a predefined field of investigation. But these commercial activities serve as an ‘anthropological vehicle’ to immerse oneself in a societal domain where the research fieldwork will unfold. As such, entrepreneurial research is a qualitative method, sharing much in common with other forms of participatory research (e.g., McNiff 2013; Hennink, Hutter, and Bailey 2020). Another aspect of entrepreneurial activity manifests itself

in our attempt to develop applications for our external partners, which in turn enhances our capacity to carry out research. We want our tools and processes to double as means for participatory observation, and to generate data for research (e.g., Siffels et al. 2022).

At the heart of our entrepreneurial activities is the development of processes, services, and products which respond to a need within the field at which our investigation is aimed. Basically, the range of services and products serve as a ‘door opener,’ allowing us to enter the domain of our research interest and to immerse ourselves in it. In their contribution to the volume *Historical Foundations of Entrepreneurship Research*, Landström and Benner (2010) explain the close relationship between entrepreneurship and the academy in terms of a development network. With reference to Blume (1985) and Funtowitz and Ravetz (1990), they state:

many research fields, especially those with major political and economic interests, are shaped by ‘development networks’ consisting of practical and academic influences, where content, direction and validation of research are outcomes of ‘negotiations’ between academic and societal interests (Landström and Benner 2010, 16).

The Data School finds itself right in the middle of such a ‘development network’ in a field with ‘major political and economic interests’: data science. By employing entrepreneurial research as our method, we can play an active part in the negotiations invoked by Landström and Benner. In an attempt not to exaggerate the impact of our research activities, we say that we are “taking part in conversations” about datafication and algorithmization, along with the development of practices for responsible AI and data science. Nevertheless, this means participating in a process of negotiating the shared understanding of, and the perspective taken with regard to, data practices and their social impact.

Because we are working with data scientists in organizations where the emerging datafied society is taking shape, we are able not only to interview people but to pick up tacit knowledge: volatile opinions and ephemeral practices that people have learned by being in a particular environment or by working with a certain tool. In the words of Michael Polanyi: “we can know more than we can tell” (Polanyi 1966, 4). To extract meaning from practices and to understand the narratives of the datafied workplace, one has to learn to use the same tools.

Similar initiatives exist elsewhere. In Hamburg, the Bureau for Social Research developed a model of carrying out student research projects

commissioned by external partners (Schnapp 2017). In these projects students can directly apply what they have learned in their studies or call into question its validity, in view of the empirical reality.

At the University of Helsinki, a research group led by Minna Ruckenstein studies developments around automation via a collaborative, open-ended approach. Using the notion of breakages as a collaborative reflection point to study algorithmic systems, the Repair project, directed by Ruckenstein, engages stakeholders to identify such breakages in AI systems (Ruckenstein et al. 2024). Researchers ask what happens after AI systems fail, promoting thinking about how they might be repaired. Whereas empirical fieldwork and stakeholder engagement are essential to this approach, this work also raises questions about the epistemological nature of breakages and repair efforts that do not receive sufficient attention (Lehtiniemi 2023).

The Critical AI & Crisis Interrogatives research group, directed by Nitin Sawhney at Aalto University, also works systematically with external stakeholders to investigate the uses and discourses of AI while developing practical solutions to ensure responsible and trustworthy AI, citizen participation, and algorithmic literacy (e.g., Gonzalez Torres and Sawhney 2023; Gonzalez Torres, Kajava, Sawhney 2023).

At the Tallinn University of Technology, Anu Masso started the Data Lab with the explicit aim of cooperating closely with external partners. This intention has taken concrete form, among other manifestations, in the publication of an 800-page book on the social aspects of generating, collecting, and processing data from the perspectives of experts. This volume, aimed at anyone working with data, featured contributions from academics and practitioners alike and was published in Estonian, and now this methodological guide to understanding a data-informed world is successfully finding its readership among Estonia's rapidly emerging technology sector. The introduction includes ten postulates – comprising the so-called Estonian Data Manifesto – that researchers and practitioners can use as starting points when working with data. It is written with explicit consideration of the Estonian context (Masso, Tiidenberg, and Siibak 2020).

And at Bremen University, the Center for Media, Communication & Information Research (ZeMKI) built two MA programs intentionally connecting with societal stakeholders and developing applicable solutions. Probably the most noted example is their app *molo news*, a platform for delivering locally relevant news (Hepp and Loosen 2019).

The above examples have several aspects in common: they are situated locally, their research practice engages with urgencies and needs identified in the societal sectors under investigation, they successfully seize

opportunities for intervention and knowledge transfer, and these projects are inherently inter- and often even transdisciplinary. Most importantly, the projects are deliberately not geared towards forms of public outreach or the use of findings to be applied but are grounded in the understanding that collaborative research with external partners advances epistemological capacity. With reference to Ruckenstein's project on breakages, the inquiry into failures opens up insights into the often invisible labor of responding to errors, compensating for undesired effects, or adapting oversight and governance. These findings – only possible through the close reading of algorithmic systems in the field – allow for better conceptualization and for intervention.

Immersing in the Field

Our attempts to immerse ourselves in the sector under investigation have many historical predecessors. These prior examples also demonstrate how these sorts of undertakings lead to the development of distinct methods and practices that not only advance research capacities but also support research subjects and perhaps even constitute change. In 1931, a group of social scientists investigated the impact of unemployment on a small Austrian community (Jahoda, Lazarsfeld, and Zeisel 1974). The village of Marienthal, known for its textile production, was struck by a sudden economic decline due to closure of the factory that was the town's main employer. The resulting widespread unemployment enabled the scholars to investigate the effects of unemployment on people and their everyday lives, including their physical and mental health. To gain more reliable insights, one of the scholars lived in the community for several weeks, and other team members also spent time in the village. Not merely observers, they developed several activities that at once provided support for the residents and served their research interests.

Providing consultation hours on health issues, organizing a clothing drive, and offering classes in sewing were means of arriving at detailed insights into the villagers' needs and their standard of living, but these efforts also built trust. The residents were likely more open to participate in interviews and spoke with greater honesty to researchers whom, they felt, not only understood the local situation but were trying to do something about it. Residents agreed to keep diaries about how they spent their days and to note down the challenges they were struggling with; they documented their spending and how they would compensate for shortfalls.

While Jehoda, Lazarsfeld, and Zeisel's social engagement in Marienthal was strongly motivated by their political attitude and their desire to support the community and ease the hardships of unemployment, it undeniably served their research efforts. With the trust they earned, they gained access and insight that allowed them to capture the social impact of unemployment. Their seminal study also demonstrates the development of novel and inventive methods to capture relevant data. Immersing themselves deeply in their subject of study, they were able to capture empirical evidence reflecting how unemployment affects all aspects of everyday life and the elements that constitute resilience or powerlessness. In a similar fashion, Gabrielle Coleman lived with San Francisco hackers and intimately mapped how their understanding of society and technology profoundly informed their software development work (Coleman 2013). The investigation of societal issues is inherently interwoven with an approach towards conceiving solutions. In their investigation of laboratory life, Steve Woolgar and Bruno Latour (2013), and in Latour's paper on the epistemic practices of geologists in the rainforest (Latour 1999), the "embedded researcher" – Latour himself – is not simply observing how knowledge develops but is, with his own expertise, contributing to this epistemic community.

The practices of the Data School can be seen in the tradition of these examples of embedded research. We enter the area under investigation and try to respond to the immediate needs that are present there, and our hosts' trust and cooperation allow us to investigate the area more closely. This relationship leads to the co-creation of applicable solutions, interventions, or educational formats, which eventually advance responsible forms of AI and data science. Akin to what Lazarsfeld and his colleagues did in Marienthal, we, too, develop specific formats that eventually help us to gather more insights. While the services and instruments we offer to our external partners help them to respond to challenges in their own organizations, they simultaneously help us gather more information.

Entrepreneurial Research in Action

Initially, as mentioned above, the Data School worked as a practicum course for students; assignments revolved around social media data analysis commissioned by external partners. The Ministry for Infrastructure and Environment requested an investigation of social media publics (mostly on the platform then known as Twitter) that were interested in and relevant to their policies. The nongovernmental organization Unicef Netherlands

had, at the time, a Twitter account with more than 200,000 followers, but it possessed little understanding of the extent to which these accounts would consist of different clusters and relate to different topics, nor did it know how to adapt the organization's communication strategy and campaigning objectives to the dynamics of Twitter. Since then we have engaged in various commissioned projects for which we employ computational methods to map and analyze public debates. These projects have explored topics such as the role of public broadcasters in these debates (Veerbeek et al. 2021; Veerbeek et al. 2022), data centers (Van Es et al. 2023), and the reporting of Covid-19 news (Nguyen and Van Es 2024). In these projects, the domain expertise of our partners played a crucial role in the interpretation of the data. Additionally, this work has contributed to academic discussions within digital methods and the Digital Humanities on the importance of tool criticism, the need to critically examine how digital tools influence knowledge production (Van Es, Schäfer and Wieringa 2018; Van Es, Wieringa and Schäfer 2021). The focus of these projects very quickly expanded to other data and to other issues than online communication.

Talks with officials from municipal safety and security departments revolved around the use of Twitter for detecting disorder, specific incidents, or emergencies during large public events. For the city of Gouda, a student team used data from more than twenty municipal datasets and the local police's data on residential burglaries to find correlations in environmental factors such as escape routes and lightning. Long-term cooperation with the city of Utrecht allowed for frequent and informal conversations with several city officials on issues of smart cities and datafication. Responsible data practice and ethics were recurring themes. Eventually, this partnership led to a joint effort to develop a process for tracing ethical issues early in a data project. The result was the Data Ethics Decision Aid (DEDA), released in 2017 (Franzke, Muis, and Schäfer 2021).

Designed as a dialogical process, DEDA brings together the various participants and managers of a municipal data project for a deliberative discussion of ethical issues. Providing DEDA workshops for municipalities allows Data School researchers to gather comprehensive information about an organization's operation and its planned data projects. The workshop process provides insights into the way data analysts, project managers and other participants are thinking about the discussed project, their responsibilities, and how their values and professional attitude affect their ethical decision-making. It also allows for estimates to be made of operational capacities and skills related to data practices (Siffels et al. 2022). Here another quality of entrepreneurial research manifests itself which

extends beyond funding research. The general dynamic became a pattern of projects at the Data School. With the European AI Act on the horizon, our understanding of public management and data ethics allowed us to develop the Fundamental Rights & Algorithms Impact Assessment (FRAIA) commissioned by the Ministry for the Interior & Kingdom Relations (Gerards et al. 2022). Subsequently, we offered a teaching module and have trained more than 250 government employees in implementing FRAIA.

Our use of digital methods and data analysis led to a collaboration with a broadcasting company involving a review of their performance metrics as well as the development of alternative metrics (Veerbeek, Van Es and Müller 2022), and it also yielded inquiries into the relationship among public discourses, radicalization, and safety issues such as death threats against city council members (Boiten et al. 2020), online antisemitism (Veerbeek et al. 2020), and correlations between street protest and social media conversations (Bakker et al. 2021). The competence developed at the Data School in applying natural language processing led to cooperation with the weekly magazine *De Groene Amsterdammer* to investigate the impact of machine learning on investigative journalism. The combined team of Data School researchers and journalists also jointly won a Dutch journalism award for one of their articles. The projects led to a co-financed PhD position for one of our junior researchers.

What started in 2013 with small student research projects emerging out of the personal network of the two founders has grown, over the years, into an interfaculty research and teaching platform, now reporting directly to the deans of the Humanities and the Science faculties. Initially, our only proposition for our external partners was some sort of data analysis, mostly network analysis of social media accounts and conversations. That focus expanded into a range of services that could be offered on different scales. We defined relevant domains where AI and data practices could be observed and which are essential to our understanding of democracy and deliberation. These domains are mostly public management and media industries. Data School now focuses on two areas of research: responsible AI and data practices, and public (political) debates (Schäfer et al. 2023). In all this time, the Data School was funded almost exclusively through its cooperation with partners outside the university. External funding allowed for significant autonomy and flexibility in project management and research capacity allocation. Time and again our research projects and teaching activities proved to have a tangible impact, shaping a track record that not only is leading to new cooperative initiatives and assignments but is also inspiring colleagues and university administrators.

Ethics of Entrepreneurial Research

Over the past years we have encountered a number of ethical issues in working with data and working in the field with external partners. The narratives concerning ‘Big Data’ or ‘AI’ shared in the corporate domain often paint a rather simplistic picture of the value of their novel possibilities or, alternately, the damage they might inflict (see Brevini 2021; boyd and Crawford 2012; Kitchin 2014a; Zwitter 2014). Both these overly optimistic and dystopian imaginations have led scholars to refer to Big Data and AI in terms of a mythology (Couldry 2014; 2017; Ziewitz 2016). Such distorted understandings sometimes result in unrealistic expectations and unethical requests on behalf of potential partners, especially when considering the fact that the Data School combines digital methods with a critical reading of societal and cultural context, as well as ethical inquiry.

A high-placed manager at a large European bank asked that we match the bank’s database of defaulted debtors with their social media contacts; apparently research had indicated that the friends of people who default on their loans are also unlikely to be able to pay off their own debts. A data warehouse, hosting transaction data from various foreign banks, inquired whether we could develop indicators for money laundering. Neither request went further than the initial conversation. The first appeared unethical to begin with, and the latter lacked the legal grounds for us to inspect the transaction data. In another instance, a branch organization of Dutch companies put us in contact with the US Department of Homeland Security. They were in the process of developing an ethical framework and wanted us to review and comment on it. We quickly decided to not go forward as we considered the entire Department of Homeland Security to be inherently unethical. The contact ceased after a conference call and a brief email exchange. A request to participate in a project commissioned by Google was also turned down due to the limiting effects of its prerequisites on academic integrity and agency. At another point we were contacted by a startup that was seeking to improve the information provided to general physicians through the analysis of large datasets. Lacking understanding of the particular data and of the health sector in general, we felt that our participation in such a research endeavor would be inappropriate. In addition, we had concerns related to the privacy-sensitive information of the startup’s data. These are just a few examples of how research integrity can be challenged due to interactions with external partners. The experience has profoundly informed our approach to selecting external partners and to defining research goals, educational objectives, and guidelines for student participation, as well as the ways we review our role as researchers.

Following certain rules and guidelines enables decision-making that is in line with ethical standards. In our practices, we strictly follow the Code of Conduct for Academic Practice (VSNU 2014; Universities of the Netherlands 2018) and, when the research topic requires it, the Association of Internet Researchers' Guidelines (Markham and Buchanan 2012; Franzke et al. 2019). Often, however, these recommendations are not sufficient in addressing the specific issues arising from cooperation with partners outside the university, working with datasets from diverse origins, and student participation. Throughout our research projects within public management, consideration of the ethics of data practices emerged as a relevant issue in itself and started to affect our research practice as much as our research agenda did (Van Schie et al. 2017). It also led to the development of our own ethical deliberation instruments, such as the Data Ethics Decision Aid for Research (Data School 2018). In addition, we influenced the university in its own process of developing policy and ethics committees for data practices. For example, as our research often relies on social media data, we called in the help of a privacy lawyer, who developed guidelines for GDPR-compliant data collection (Gerritsen 2021). It became an official university guideline.

Checks and balances for research integrity manifest themselves on procedural, financial, and informal levels. The procedural level consists of the selection process of the partners and the terms of collaboration included in the cooperation agreement. On the financial level, budget planning ensures that we are never dependent on a single partner. Together, the contract and our financial planning allow us the option of walking away from any project that would seem to compromise our integrity. The informal level is essential, too: we have agreed that if any member of the Data School management team should feel uncomfortable, for any reason, with a potential external partner, a veto can prevent the partnership from going forward.

Funding from external partners expanded our capacities to hire researchers, extend contracts, and expand teams as needed. Importantly, the projects that we conduct are rather small-scale, so we are never reliant on a single partner. Moreover, the university offers us significant basic support by providing us with office space as well as administrative and legal support within their capacity to do so; the academic staff are on the university payroll.

Setting Goals and Basic Values

From experience, we learned what to look for in an external partner and how to effectively communicate and negotiate our values. The entrepreneurial

activities we engage in are guided by our initial research interest. Within these topics, we develop collaborative research projects and educational activities. Our cooperation with external partners must present the prospect of advancing our insights into the transformation of organizations through datafication, and it should also create valuable learning and career opportunities for students. Lastly, it should offer the possibility of intervention. We formulated this sense of our values in the form of three prerequisites for our activities:

- We need to be where the societal and technological change manifests itself.
- Research and teaching must be intrinsically connected.
- Our activities must yield insights relevant for academic research and make an impact.

We consider these aspects before embarking on any sort of cooperation with external partners. Lisa Spiro has proposed several values that she considers essential for the digital humanities (Spiro 2012): openness, collaboration, collegiality and connectedness, diversity, and experiment. These values largely overlap with those common to academic institutions. However, this list can serve as a useful set of guidelines for dealing with external partners (and for shaping the work-related ethics of your team). Because external partners naturally follow agendas that are somewhat different from an academic agenda, these values constitute helpful guidelines to explain the limits of cooperation and, more importantly, to define requirements for collaboration. Referring to our objectives, values such as openness and collaboration, as well as collegiality and connectedness, are essential. External partners, therefore, have to be open to providing access to corporate information and technical support if needed and applicable, but most importantly, they need to offer the time to answer queries about the external partner's perspectives on data practices. Collegiality is at the core of this aspect, and it is a quality that requires all participants – including academic staff, practitioners, and students – to create an atmosphere that encourages mutual cooperation, support for one another, the solving of problems together, and joint efforts directed towards realizing shared objectives. According to the terms of contract research at our university, research activities are centered around best efforts rather than results.

The value of openness plays out on several levels. First, because the research projects are inherently interdisciplinary and sometimes even transdisciplinary, openness to different methodological approaches,

values, and objectives must be embraced by all participants. It takes time to overcome the inherent differences of specific epistemic cultures, to find a common language, and to integrate the different forms of explicit and tacit experience and knowledge. Second, openness also relates to access to the external partners' infrastructure and data resources, as well as the sharing of research findings. This is at times difficult, as partners might not want to share sensitive findings that can make their way to competitors, and findings might also be subject to security concerns. In such instances, the researchers and the external partner try to come to terms by either delaying publication through an embargo or by excluding sensitive information when a text is prepared for publication. Third, there should be openness considering the varying motivations, agencies, and codes of conduct of the different participants, external partner organizations and their employees, university staff, and students. At the Data School, we always emphasize that we are action researchers, not activist researchers. We are also not investigative journalists. With reference to Marshall McLuhan, we commit to a "technique of suspended judgement", which McLuhan – drawing from Bertrand Russell, and A. N. Whitehead – described as a practice of discovery (1964, 68). For us, this can sometimes mean holding back critical commentary or refraining from dismissive judgement, which otherwise might prematurely end conversations that we need to continue in order to gain insight or to intervene. But it also means granting others the space to voice their perspective, to share their insights and knowledge, to accept and to embrace their subject expertise and practice experience as relevant for the process of knowledge generation.

Selecting Partners

Most parties interested in cooperation learn about our services through publications in professional magazines or appearances at practitioners' conferences, business fairs, or expert meetings, though we increasingly get referrals from recent graduates or professionals taking part in our education for professionals programs. Figure 4.1 shows, schematically, the selection process of partners for collaboration: an initial meeting, follow-up meeting(s), planning collaboration, and follow-up. After initial contact, a brief phone call is scheduled to determine whether there is enough common ground for an in-person meeting. This initial contact lays out basic parameters of cooperation needed for a common research interest, the acceptance of academic freedom, and information about financial contribution.

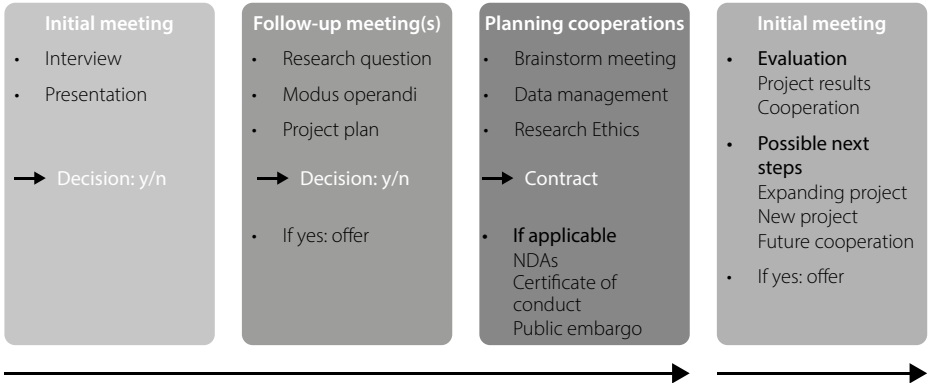


Figure 4.1. Schematic selection process.

In a follow-up meeting we explain our methods and modus operandi and present examples of our work. The external partner reports their data issue or research question and elaborates on their perspective towards datafication. The challenges and opportunities of potential cooperation are discussed. Here we point the external partner to our objectives (such as learning opportunities for students), our research agenda concerning datafication, publication opportunities, and our university guidelines concerning best effort rather than results.

Afterwards, both parties deliberate whether they are interested in developing a research project together. This discussion allows us to reflect on whether our objectives and values overlap sufficiently with those of the external party. If both parties agree, the follow-up meeting focuses on developing a research project. At this stage irreconcilable differences may well manifest themselves, or perhaps a joint effort will appear to be not feasible or, for whatever reason, to be undesirable. But if that meeting does produce a viable research project, then a third meeting is scheduled. There we sketch the *modus operandi* of cooperation; define common guidelines; speak about data management, data ethics, and non-disclosure agreements if applicable and necessary; and anticipate the resources that will be needed by both parties. A contract is subsequently drafted and signed.

Acquisition and managing the cooperation is a time-consuming activity that is difficult to delegate to designated impact officers. The expertise of the researchers, and their social capabilities to build and maintain networks outside the university, are essential for collaborative research projects. We developed activities such as network events, expert meetings, and seminars where we would meet our partners and other parties who might be interested in collaboration. However, this particular activity is a form of invisible labor,

work that is hardly recognized or understood, let alone rewarded, within and by universities. On the positive side, we developed many long-term collaborations and have carried out multiple projects.

Knowledge Dissemination

In entrepreneurial research, as explained, the researcher has to engage simultaneously with different communities. The primary reference group is certainly the peer groups where dissemination manifests itself in peer-reviewed publications, conferences, the supervision of PhD students, and the teaching and development of appropriate courses. Given the interdisciplinary quality of the research projects, this is already challenging. In addition, dissemination is supposed to focus on the societal sector in which the research takes place. Here, we publish in professional magazines and present research for the practitioners. Other platforms, such as practitioners' conferences, as well as participation in workshops and expert meetings and availability as speakers or conversation partners, provide opportunities to disseminate research, find collaborators, solicit feedback, and, most importantly, engage in critical conversations on the use of data and AI.

Reaching out to journalists covering the societal sector or topics relevant to the research interest is highly recommended. Even if such outreach does not directly lead to citations or media coverage of the individual researcher, it fosters the entrepreneurial scholar being regarded as an expert on the topic as well as a person who is approachable for questions. Members of our research group had numerous, often hour-long conversations with journalists from a wide range of publications to provide background information for an article the journalist was working on or just to have an informed discussion on data practices for journalists and on the impact of datafication in general.

Challenges of Entrepreneurial Research

The approach to research sketched above provides many opportunities. It is a creative way of seizing opportunities for research that takes place within particular societal sectors. It opens up many ways for intervention and change, for learning and knowledge transfer, as well as for funding. These opportunities, described at length above, are summarized in the table below. When engaging with entrepreneurial research one also encounters a

number of challenges. These challenges, also listed in the table below, will be addressed in detail in the following section.

Table 4.1. Benefits and Challenges

Benefits	Challenges
Connecting with the knowledge and expertise of professionals	Shared vocabulary (between disciplines and in translation to societal partner)
Unique access to internal discourses and data	Different speeds between academia and society
Providing learning opportunities for students/recent graduates	Protecting academic integrity
Independence through external funding	Types of (applied) questions and fragmentation of academics' research agendas
Theory meets practice = empirical grounding	Institutional infrastructures, e.g., administrative, legal, technological
Enriching fundamental research	Recognition and rewards

Shared vocabulary

Often researchers find it difficult to work in an interdisciplinary fashion because doing so requires them to acquaint themselves with the practices, methods, jargon, and publication culture of a different discipline. It becomes even more difficult if research projects include partners outside the university, each having their own practices, methods, jargon, and organizational culture. In addition to reading papers from the cooperating disciplines, researchers also have to become acquainted with the knowledge present and disseminated within the relevant societal sector, as expressed concretely in reports or more ephemeral formats such as memos, notes, etc.

Different speeds

Academic organizations operate at a different speed and according to different parameters than many organizations from other societal sectors. These differences inevitably give rise to several practical challenges for entrepreneurial research. The infrastructure and expertise needed to support third-revenue projects in setting up contracts, carrying out project management, providing legal support, etc. are often not developed in academic institutions. Traditionally, research projects revolve around the budgeting and duration of the PhD thesis, which is usually scheduled for periods ranging from three to six years. Most of the research projects carried

out in the fashion of entrepreneurial research are shorter than a year; they often last several months. The response time of academic administrative units and researchers is often slower than external partners might expect. Decision-making and project development, but also acquisition trajectories and contracting, can suffer or even lead to termination because of such disparities. Additionally, attention in the public realm is determined by the fast pace of media publications, which affects the dissemination of research results and the decision of when and where to publish them.

Academic integrity

As we are not mere observers but active participants, we as scholars must develop antennae to properly understand how our research objectives might clash with the agenda of the cooperating party, and to what extent such varying priorities might compromise our integrity. Our research practices are not limited to observation: we are participating, we are actively engaging and intervening. We are not onlookers or bystanders but are invested collaborators; and sometimes we might be even complicit in undesirable ways. One of our projects led to the termination of our corporate liaison's contract by his employer. The project carried out by our students revealed that this employee had failed to stimulate necessary cooperation from other organizational units and departments.

The need for an awareness of ethical issues and responsible data practices, as well as the inherently political quality of data practices, became apparent to us in almost every research project that was undertaken. These concerns are now explicitly addressed in our research mission and our objectives; before starting a project, we always reflect on how it advances our research agenda, aligns with our values, and complies with codes of conduct and regulations. Signaling to external parties that they must critically assess their own data practices and related policies also helps us to prevent actions or incidents which might compromise our own ethical standards. Ethics and responsible data practices can be made explicit through the implementation of an ethics review as part of every project. Using our own Data Ethics Decision Aid (DEDA), a process initially developed with the city of Utrecht as an external partner, students and external partners structurally review a data project and identify possible ethical issues.

External funding allows researchers to enjoy an exceptionally autonomous position, provides flexibility in executing research and getting work done, and creates independence from the often neglected power asymmetries and dependencies within the university and instituted through traditional

research funding. Nevertheless, external funding always raises issues of academic integrity. Here our *modus operandi*, with clear procedural rules separating budgetary flexibility and the possibility that management team members can veto – without further argument – any project they do not feel comfortable with, has allowed us to remain largely independent and able to walk away from any project that goes against our values.

Rewards and recognition

Other risks of entrepreneurial research entail marginalization, time issues, and institutional restraints. Developing entrepreneurial research requires activities that lie outside faculty duties: developing and maintaining an extensive network within the sector where the research takes place; developing services and products that will eventually facilitate collaboration, participatory observation, data collection, and funding; promoting those services and products; finding external partners and forming collaborations with them; optimizing results and implementing them in societal sectors; managing a team and securing needed funding; dealing with a significant amount of additional administrative work arising from cooperation agreements and contracts, licenses, contracts, and other legal aspects; and thinking of branding and promoting educational formats among students and professionals. In fact, entrepreneurial researchers have two jobs: one as entrepreneur and the other as researcher. At times the burden of carrying out both tasks and bringing together these two worlds can be exceptionally difficult. However, we experience the perpetual oscillation between academia and societal sectors as an inspiring and productive experience.

Within academia, neoliberal policies have emphasized the immediate application of research results. Impact, not only through peer-reviewed publication but also through dissemination to non-academic audiences and application to problem-solving or policy development in various societal sectors, has been added to the list of academic responsibilities. However, we have experienced the university as an institution widely unprepared for the challenges of structurally interacting with societal sectors. There are hardly any incentives for engaging in the range of activities listed above. Scholars on temporary contracts in particular cannot risk their careers in the service of developing an entrepreneurial research agenda, which has to be developed with limited support from the academic host institution.

But the tide is changing. University policymakers are not only calling for change but are initiating it (Kummeling et al. 2023). By signing the DORA declaration, thousands of academic organizations around the globe have

committed to developing better ways of measuring and evaluating academic performance. In 2023, Utrecht University withdrew from the Times Higher Education World University Rankings, arguing not only that its analytical methods were insufficient but, most importantly, that the rankings process contradicts their core value of academic collaboration.

The Royal Netherlands Academy for the Arts and Sciences lists, as key performance indicators for accreditation evaluation, the development of ‘academic’ and ‘societal products’, the dissemination to non-academic audiences, and the impact of research findings in policymaking and problem-solving. We hope that the universities will move from mere lip-service towards promoting impact, interdisciplinary cooperation, and societal impact and actually support and reward these activities. For now, we still note that the institutional infrastructures are “slow and reluctant to adapt to innovations in creative research approaches” (Cornish et al. 2023).

Conclusion

To reiterate: entrepreneurial research is a creative process of finding opportunities to carry out research and to apply findings. Entrepreneurial researchers seek opportunities to immerse themselves in a societal sector in order to study it or a phenomenon unfolding there. Like action researchers, they also want to provide applicable findings, to intervene, and to disseminate learnings effectively. However, their core motivation is to gain better insights from within, to understand a phenomenon – in our case, how AI and Big Data change democracy – through studying it up close. The entrepreneurial activities provide us with the opportunity to get close to the phenomenon, understand it within its complex socio-technical context, and to devise appropriate means for learning and change. Entrepreneurial research is about empirically understanding a phenomenon and seizing opportunities to advance both societal change and academic research. To do so, entrepreneurial research avails itself of a range of activities, of which some might be commercially oriented and aimed at funding, others at creating and sustaining knowledge transfer and learning, or at developing applicable solutions to respond to problems. Many activities serve the goal of immersion in the field of investigation and gain access to rich data that would not otherwise have been accessible. The activities are merely a means of conducting research. The rewards of this approach are plentiful: they range from accessing rich data and gaining deep insights into developing additional expertise and novel education formats, to career

opportunities for students, the application of findings, and societal impact. However, this approach raises points of concern and requires tactics, skills and actions that should be discussed and revisited on the level of research associations and university policymakers. As universities put greater emphasis on societal impact and community-engaged learning, and as they are now opening up their processes of generating knowledge through inter- and even transdisciplinary collaboration, they must address the epistemological and administrative challenges brought forth by these approaches.

Acknowledgements:

This chapter is the result of a collective effort to shape the distinct practices and methods at Utrecht University's Data School. Neither this chapter nor the many projects that we have concluded over the years would have been possible if it was not for the team effort. All team member at the Data School contributed to developing our understanding of entrepreneurial research and shaped the activities, products, services, and tools which advance our research. We are indebted to our external partners for their trust and for providing access to their organizations and data, as well as for the learning opportunities given to our students, and their willingness to apply our findings.

References

- Bakker, Jeroen, Daphne Cornelisse, Sahra Mohamed, Mirko Tobias Schäfer and Joris Veerbeek. 2021. "Van Scherm naar Straat. Hoe Sociale Media-Conversaties Protest op Straat Mobiliseren." Research report. Utrecht Data School. Utrecht University.
- Barton, Allen H. 1979. "Paul Lazarsfeld and Applied Social Research: Invention of the University Applied Social Research Institute." *Social Science History* 3 (3-4): 4-44.
- Berry, David M. 2012. *Understanding Digital Humanities*. Houndmills, Basingstoke and Hampshire: Palgrave Macmillan.
- Blume, Stuart. 1985. "After the Darkest Hour: Integrity and Engagement in the Development of University Research." In *The University Research System: The Public Policies of the Home of Scientists*, edited by B. Wittrock and A. Elzinga, 139-163. Stockholm, Sweden: Almqvist & Wiksell International.

- Boiten, Max, Maica Hopstake, Marjolein Krijgsman, Roelof Muis, Peter Mullaart, Sander Prins, Mirko Tobias Schäfer, Joris Veerbeek. 2020. "Online Bedreigingen in het Lokaal Bestuur." In *Aard en Omvang van Dader- en Slachtofferschap van Cyber- en Gedigitaliseerde Criminaliteit in Nederland*, edited by M. G. C. J. Beerthuizen, T. Sipma, and A. M. van der Laan, vol. Cahier 2020–15, 46–58. The Hague: WODC.
- boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15 (5): 662–679.
- Brevini, Benedetta. 2021. "Creating the Technological Saviour: Discourses on AI in Europe and the Legitimation of Super Capitalism." In *AI for Everyone?: Critical Perspectives*, edited by Pieter Verdegem, 145–160. London: University of Westminster Press.
- CA-ICT, Nederland ICT, and CIO Platform Nederland. 2017. "Arbeidsmarktonderzoek Topsectoren 2017." <https://www.caict.nl/arbeidsmarktonderzoek-topsectoren/>.
- Coleman, E. Gabriella. 2013. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton, NJ: Princeton University Press.
- Couldry, Nick. 2014. "Inaugural: A Necessary Disenchantment: Myth, Agency and Injustice in a Digital World." *Sociological Review* 62 (4): 880–897. <https://doi.org/10.1111/1467-954X.12158>.
- Couldry, Nick. 2017. "The Myth of Big Data." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 235–240. Amsterdam: Amsterdam University Press.
- Dijk, José van, Thomas Poell, and Martijn de Waal. 2016. *De Platform Samenleving: Strijd Om Publieke Waarden in Een Online Wereld*. Amsterdam: Amsterdam University Press.
- Dooley, Lawrence, and David Kirk. 2007. "University-industry Collaboration: Grafting the Entrepreneurial Paradigm onto Academic Structures." *European Journal of Innovation Management* 10 (3): 316–332. <https://doi.org/10.1108/14601060710776734>.
- Etzkowitz, Henry. 2003. "Research Groups as 'Quasi-Firms': The Invention of the Entrepreneurial University." *Research Policy* 32 (1): 109–121.
- Etzkowitz, Henry. 2014. "Making a Humanities Town: Knowledge-Infused Clusters, Civic Entrepreneurship and Civil Society in Local Innovation Systems." *Triple Helix* 2 (December): 1. <https://doi.org/10.1186/s40604-014-0012-z>.
- Franzke, Aline, Anja Bechnman, Michael Zimmer, and Charles Ess. 2019. "Internet research: Ethical guidelines 3.0." Association of Internet Researchers. <https://members.aoir.org/resources/Documents/IRE%203.0%20-%20final%20distribution%20copy%20of%20AoIR%20members.pdf>.
- Funtowicz, Silvio O., and Jerome R. Ravetz. 1990. *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer Academic Publishers.

- Gonzalez Torres, Ana Paula, and Nitin Sawhney. 2023. "Role of Regulatory Sandboxes and MLOps for AI-Enabled Public Sector Services." *The Review of Socionetwork Strategies* 17: 297–318. <https://doi.org/10.1007/s12626-023-00146-y>.
- Gonzalez Torres, Ana Paula, Kaisla Kajava, and Nitin Sawhney. 2023. "Emerging AI Discourses and Policies in the EU: Implications for Evolving AI Governance." In *Artificial Intelligence Research*, edited by Anban Pillay, Edgar Jembere, and Aurore J. Gerber. SACAIR 2023. Communications in Computer and Information Science, vol. 1976. Cham: Springer. https://doi.org/10.1007/978-3-031-49002-6_1
- Goulet, Jean-Guy A., and Bruce Granville Miller. 2007. "Embodied Knowledge: Steps toward a Radical Anthropology of Cross-Cultural Encounters." In *Extraordinary Anthropology: Transformations in the Field*, edited by Jean-Guy A. Goulet and Bruce Granville Miller, 1–13. Lincoln, NE: University of Nebraska Press.
- Grant, Paul, and Lew Perren. 2002. "Small Business and Entrepreneurial Research: Meta-Theories, Paradigms and Prejudices." *International Small Business Journal* 20 (2): 185–211.
- Harman, Grant. 2005. "Australian Social Scientists and Transition to a More Commercial University Environment." *Higher Education Research & Development* 24 (1): 79–94. <https://doi.org/10.1080/0729436052000318587>.
- Hepp, Andreas, and Wiebke Loosen. 2019. "Molo. News: Experimentally Developing a Relational Platform for Local Journalism." *Media and Communication* 7 (4): 56–67.
- Jahoda, Marie, Paul F. Lazarsfeld, and Hans Zeisel. 1974. *Marienthal: The Sociography of an Unemployed Community*. New Brunswick, NJ: Transaction Publishers.
- Kitchin, Rob. 2014a. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1 (1): 1–12.
- Kitchin, Rob. 2014b. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
- Landström, Hans, and Mats Benner. 2010. "Entrepreneurship Research: A History of Scholarly Migration." In *Historical Foundations of Entrepreneurship Research*, edited by Hans Landström and Franz Lohrke, 15–45. Cheltenham: Edward Elgar Publishing.
- Landström, Hans, and Franz Lohrke. 2010. *Historical Foundations of Entrepreneurship Research*. Cheltenham: Edward Elgar Publishing.
- Latour, Bruno. 1999. "Circulating Reference: Sampling the Soil in the Amazon Forest." In *Pandora's Hope: Essays on the Reality of Science Studies*, 24–79. Cambridge, MA: Harvard University Press.
- Latour, Bruno, and Steve Woolgar. 2013. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.
- Lee, Yong S. 1996. "'Technology Transfer' and the Research University: A Search for the Boundaries of University-Industry Collaboration." *Research Policy* 25 (6): 843–863. [https://doi.org/10.1016/0048-7333\(95\)00857-8](https://doi.org/10.1016/0048-7333(95)00857-8).

- Lee, Yong S. 2000. "The Sustainability of University-Industry Research Collaboration: An Empirical Assessment." *The Journal of Technology Transfer* 25 (2): 111–133.
- Lehtiniemi, Tuukka. 2023. "Contextual Social Valences for Artificial Intelligence: Anticipation that Matters in Social Work." *Information, Communication & Society*: 1–16.
- Manovich, Lev. 2011. "Trending: The Promises and the Challenges of Big Social Data." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 460–475. Minneapolis, MI: University of Minnesota Press.
- Markham, Annette, and Elizabeth Buchanan. 2012. "Ethical Decision-Making and Internet Research: Recommendations from the AOIR Ethics Working Committee (Version 2.0)." Association of Internet Research. <http://aoir.org/reports/ethics2.pdf>.
- Masso, Anu, Katrin Tiidenberg, and Andra Siibak. 2020. "Kuidas mõista andmestunud maailma? Metodoloogiline teejuht." TLÜ Kirjastus. <https://www.tlu.ee/pood/home/325-kuidas-moista-andmestunud-maailma-metodoloogiline-teejuht.html>.
- Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York, NY: Houghton Mifflin Harcourt.
- Ministerie van Economische Zaken, and Ministerie van Onderwijs, Cultuur en Wetenschap. 2013. *Nederlandse Oplossingen Voor Wereldwijde Problemen*. Rotterdam: Vijfkeerblauw. <https://www.rvo.nl/sites/default/files/2013/11/Nederlandse-oplossingen-voor-wereldwijde-uitdagingen-Topsectoren-en-Horizon-2020.pdf>.
- Nelson, Andrew J. 2005. "Cacophony or Harmony? Multivocal Logics and Technology Licensing by the Stanford University Department of Music." *Industrial and Corporate Change* 14 (1): 93–118.
- Nguyen, Dennis and Karin van Es. 2024 [Forthcoming]. "Exploring the Value of Computational Methods for Journalistic Self-Reflection and 'Accountability.'" *Journalism Studies*.
- Pattnaik, Pinaki Nandan, and Satyendra C. Pandey. 2014. "University Spinoffs: What, Why, and How?" *Technology Innovation Management Review* 4 (12): 44–50.
- Perren, Lew, and Monder Ram. 2004. "Case-Study Method in Small Business and Entrepreneurial Research: Mapping Boundaries and Perspectives." *International Small Business Journal* 22 (1): 83–101.
- Pilegaard, Morten, Peter W. Moroz, and Helle Neergaard. 2010. "An Auto-Ethnographic Perspective on Academic Entrepreneurship: Implications for Research in the Social Sciences and Humanities." *The Academy of Management Perspectives* 24 (1): 46–61.
- Polanyi, Michael. 1966. *The Tacit Dimension*. Chicago, IL: University of Chicago Press.
- Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr. 1996. "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning

- in Biotechnology." *Administrative Science Quarterly* 41 (1): 116–45. <https://doi.org/10.2307/2393988>.
- Richterich, Annika. 2018. *The Big Data Agenda: Data Ethics and Critical Data Studies*. London: University of Westminster Press.
- Rogers, Richard. 2013. *Digital Methods*. Cambridge, MA: MIT Press.
- Ruckenstein, Minna, Sarah Pink, Melisa Duque, Blanca Callén. 2024. "Anticipatory Repair: Reframing Breakage through a Futures Agenda." Interview with Minna Ruckenstein and Sarah Pink. In *Disena* 24, January 24. <https://doi.org/10.7764/disena.24>.
- Schäfer, Mirko Tobias, and Karin van Es. 2017. "New Brave World." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 13–22. Amsterdam: Amsterdam University Press.
- Schnapp, Kai-Uwe. 2017. "Das Hamburger 'Projektbüro Angewandte Sozialforschung.'" *Zeitschrift für Politikwissenschaft* 27 (4): 529–547.
- Shane, Scott Andrew. 2004. *Academic Entrepreneurship: University Spinoffs and Wealth Creation*. Cheltenham: Edward Elgar Publishing.
- Spiro, Lisa. 2012. "'This Is Why We Fight': Defining the Values of the Digital Humanities." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 16–35. Minneapolis, MN: University of Minnesota Press.
- Steinbuch, Maarten. 2016. "Towards the 4th Generation University." <https://steinbuch.wordpress.com/2016/07/23/towards-the-4th-generation-university/>.
- Stuart, Toby E. 2000. "Interorganizational Alliances and the Performance of Firms: A Study of Growth and Innovation Rates in a High-Technology Industry." *Strategic Management Journal* 21 (8): 791–811.
- Ucbasaran, Deniz, Paul Westhead, and Mike Wright. 2001. "The Focus of Entrepreneurial Research: Contextual and Process Issues." *Entrepreneurship Theory and Practice* 25 (4): 57–80.
- Van Dijck, José, and Thomas Poell. 2013. "Understanding Social Media Logic." *Media and Communication* 1 (1): 2–14.
- Van Es, Karin, Maranke Wieringa and Mirko Tobias Schäfer. 2018. "Tool Criticism: From Digital Methods to Digital Methodology." ACM WS.2 2018 Proceedings of the 2nd International Conference on Web Studies. <https://doi.org/10.1145/3240431.3240436>.
- Van Es, Karin, Mirko Tobias Schäfer and Maranke Wieringa. 2021. "Tool Criticism and the Computational Turn: A 'Methodological Moment' in Media and Communication Studies." *M&K Medien & Kommunikationswissenschaft* 69 (1): 46–64.
- Van Schie, Gerwin, Irene Westra, and Mirko Tobias Schäfer. 2017. "Get Your Hands Dirty: Doing Ethical Research in a Big Data Era." In *The Datafied Society: Social Research in the Age of Big Data*, edited by Mirko Tobias Schäfer and Karin van Es, 183–200. Amsterdam: Amsterdam University Press.

- Veerbeek, Joris, Karin van Es and Eggo Müller. 2021. "Professional follow-up communication as an indicator of public value: Analyzing Dutch PSB's role in societal debates through topic modelling of press coverage." *Journal of Communication Inquiry*. <https://doi.org/10.1177/01968599211039213>.
- Veerbeek, Joris, Karin van Es and Eggo Müller. 2022. "Public Broadcasting and Topic Diversity in The Netherlands: Mentions of Public Broadcasters' Programming in Newspapers as Indicators of Pluralism." *Javnost*. <https://doi.org/10.1080/13183222.2022.2067956>.
- Veerbeek, Joris, Jeroen Bakker, Sahra Mohamed, Meilina Reksoprodjo, Tom Stienen, Mirko Tobias Schäfer. 2022. Online antisemitisme [Online antisemitism]. Research report, Utrecht University. <https://dataschool.nl/en/nieuws/antisemitisme-in-de-nederlandse-online-sfeer-in-2020/>.
- Voorn, Eduard. 2017. "Professor Emile Aarts: 'Heb Als Ondernemer de Guts Om Big Data-Stappen Te Zetten.'" *De Ondernemer*, May 9, 2017. <https://www.deondernemer.nl/nieuwsbericht/171036/professor-emile-aarts-heb-als-ondernemer-de-guts-om-big-data-stappen-te-zetten>.
- VSNU. 2014. *Code of Conduct for Academic Practice: Principles of Good Academic Teaching and Research*. Translated by Metamorfose Vertalingen BV. The Hague: Association of Universities in The Netherlands. [http://www.vsnul.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_\(version2014\).pdf](http://www.vsnul.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_(version2014).pdf).
- Wright, Mike. 2007. *Academic Entrepreneurship in Europe*. Cheltenham: Edward Elgar Publishing.
- Ziewitz, Malte. 2016. "Governing Algorithms: Myth, Mess, and Methods." *Science, Technology, & Human Values* 41 (1): 3–16. <https://doi.org/10.1177/0162243915608948>.
- Zwitter, Andrej. 2014. "Big Data Ethics." *Big Data & Society* 1 (2): 1–6.

About the Author

Mirko Tobias Schäfer is Associate Professor of AI, Data & Society at Utrecht University's research area *Governing the Digital Society* and the Department for Information & Computing Sciences. Mirko is co-founder and Sciences lead of the Data School. He studies the datafication of public management and engages in the development of responsible and accountable AI and data practices.

5. Open Government Partnership (OGP): Balancing Expertise, Practice, and the Academy

Mary Francoli and Daniel J. Paré

Abstract

In this chapter, the authors reflect upon the institutional challenges and professional benefits they have encountered as a consequence of their data-focused policy work and research for, and engagements with, the Open Government Partnership (OGP). Both authors have extensive experience working with OGP's Independent Reporting Mechanism (IRM). Drawing from this experience, they set out three ways in which data-focused policy work often finds itself marginalized in relation to more 'typical' academic research, and contrast these with what they identify as five benefits arising from hands-on applied research. They wrap up their discussion by suggesting steps that can be taken to move toward more fully recognizing the value and merit of such scholarly endeavors.

Keywords: Policy research; Applied research; Transdisciplinary networking; Societal impact

OGP: A brief history

The Open Government Partnership (OGP) is an international body that champions open data and open government as tools for improving transparent, accountable, and participatory governance. OGP's founding can be traced back to former US president Barack Obama's 2010 introduction to the UN General Assembly of a multilateral initiative seeking to reinforce open government in the US and worldwide that pledged to promote transparency

and civic participation, and to tackle corruption through data-based initiatives (The White House, 2010). The following year representatives from eight countries – Brazil, Indonesia, Mexico, Norway, the Philippines, South Africa, the United Kingdom, and the United States – convened in New York City with members of nine civil society organizations to officially launch OGP. Formally established in 2011, OGP was founded on three pillars (Open Government Declaration 2011): to increase the availability of information about governmental activities; to implement the highest standards of professional integrity throughout our administrations; and to increase access to new technologies for openness and accountability.

As of January 2024, OGP's membership comprises seventy-five countries and 104 local governments. Each member government commits itself to meeting specific criteria and participation requirements in areas of fiscal transparency, access to information, citizens' participation, and public disclosure. In implementing their respective OGP initiatives, member governments follow similar approaches to designing national or local action plans, amending and/or implementing various laws and regulations, opening new communication channels and practices with civil society representatives and other interested stakeholders, and harnessing new technologies to enhance the availability and accessibility of government data. To date, the commitment to these initiatives has varied widely across the participating governments, as have the results. Some have claimed successes in particular policy areas (e.g., open contracting in Ukraine and beneficial ownership in the UK), whereas others have withdrawn completely (e.g., Hungary [2012–2016], Luxembourg [2016–2023], Turkey [2012–2017]), and a few have been suspended for failing to meet their OGP obligations (e.g., Azerbaijan [2011–2023], El Salvador [2011–2023], Pakistan [2016–2022], Tanzania [2011–2017], Trinidad and Tobago [2012–2019]).

Canada formally joined OGP in 2012. Since then, it has implemented four national action plans (2012–2014, 2014–2016, 2016–2018, 2018–2021) and is currently carrying out its fifth plan. Each plan is designed and enacted in consultation with the federal government, civil society groups, and other stakeholders. As part of an OGP-led pilot experiment, the province of Ontario designed and implemented a sub-national action plan in 2016–2017. In September 2021, it launched a second action plan, focusing on the building of a framework for trustworthy AI use in the province. In 2020, the province of Quebec became a local member of OGP. Its 2021–2023 action plan has five commitments structured around four themes: open data, open science, digital public participation, and open software.

The focus of the Canadian government's action plans has shifted over the past decade from an initial concentration on enhancing transparency and openness to, more recently, pursuing principles of participation and collaboration. Its first three action plans, designed and implemented between 2012 and 2018, centered around openness-related objectives, including open information, open data, open dialogue, open by default, and fiscal transparency. The fourth plan (2018–2020) was anchored by principles of inclusion, participation, and impact. Canada's fifth national action plan (2022–2024) is structured around five themes: climate change and sustainable growth; democracy and civic space; fiscal, financial and corporate transparency; justice; and open data for results.

The discussion in the remainder of this chapter details the institutional challenges and professional benefits we have encountered as a result of our data-focused policy work for and engagements with OGP.

Applied OGP-Focused Scholarship: The Institutional Challenges

We have both worked with OGP's Independent Reporting Mechanism (IRM), an independent body that assesses the progress of member governments towards completing their open government commitments and fulfilling of OGP processes and vision. Mary Francoli has been involved with OGP since 2013. From 2013 to 2017 she was the IRM researcher for Canada, a role in which she assessed the Canadian federal government's performance in meeting the commitments spelled out in its first and second national action plans on open government (Francoli 2015; 2016; 2017). In 2017, she was appointed to the IRM's International Expert's Panel (IEP). She held this role until 2024 and chaired the group twice. The IEP directly oversees the work of the IRM and plays a key role in ensuring the quality and independence of IRM reports. It also is instrumental in setting the vision for the IRM, reviewing the methodology used for assessments, and communicating data and findings generated by IRM research. Daniel J. Paré has been involved with the IRM since 2017, when he conducted an independent assessment of the Ontario government's progress towards its 2016–2017 action plan as part of OGP's sub-national pilot experiment (later to become the OGP Local Program) (Paré 2018). In 2019, he became the IRM researcher for Canada, assessing the Canadian federal government's performance in relation to its fourth national action plan on open government (Paré 2021). He is currently working on assessing the government's performance in relation to Canada's fifth plan. After detailing the benefits we associate with our multi-domain

spanning OGP-related work, we offer some suggestions aimed at improving the recognizability of such efforts within academia.

When engaging in ‘hands-on’ research of the type outlined above, one must frequently undertake a complex negotiation of competing priorities. At issue is a disconnect between high-level institutional rhetoric about valorizing collaborative research with positive societal impacts versus the career-related consequences frequently awaiting those who engage in time-sensitive, ground-level, policy-relevant research that speaks to both academic and practitioner audiences and whose impacts are observable foremost in the form of changes in organizational policy and/or practice as opposed to being found in journal- and author-level publication metrics.

There are three ways this type of scholarship finds itself marginalized *vis-à-vis* more ‘typical’ academic research. The first pertains to the standing of this type of work within academic disciplines. Specifically, the applied data- and policy-centric priorities of the IRM’s research, analysis, and evaluation are often perceived as falling outside the intellectual priorities of individual disciplines. This, though, is hardly a new phenomenon. In the field of media and communication studies, for instance, the peripheralization of policy-oriented work on the grounds that “policy priorities are not the same as intellectual priorities” (Carey 1978, 116) or because policy researchers “cannot be expected to make contributions to the general field of knowledge of any specific academic discipline” (Lazarsfeld 1975, 211) has been manifest for at least a half-century.

The second factor contributing to the marginalizing of more practice-oriented research pertains to the principal outputs generated by this type of work. Here, there are two complementary considerations. The first pertains to the outputs themselves. Contrary to what occurs within the typical academic assessment structure, the principal products of IRM work are published in the form of assessment reports. Second, the primary audience for these outputs is not other scholars but rather members of government, civil society, and other interested stakeholders, academic or otherwise. As such, these research outputs and their relevance do not fall neatly into the standard categories most valorized by workplace evaluation and promotion committees, or by granting councils. Work such as that done as part of the IEP is even more difficult to have recognized within academic institutions precisely because such efforts do not lead to the production of written outputs in the same ways that work by IRM researchers does. The best that scholars participating in the IEP can hope for on the institutional front is that their efforts *might* be recognized as a form of service to the profession.

An additional source of consternation in the face of this peripheralization is that IRM research is, importantly, peer reviewed and has direct, real-world impacts. In terms of peer review, IRM research is subjected to high-level, quasi-anonymous quality control that is intended to ensure methodological rigor and accuracy of reported findings. Reports produced by IRM researchers are reviewed multiple times by IRM staff and by members of a reviewer pool comprised of subject matter experts, multiple civil society members, and several government representatives. They also are open to public scrutiny and comment. Quality of the reviewing process is overseen by the IEP and is evaluated regularly. Here, too, we are not dealing with a new debate. The academic reward structure places a premium on publishing in peer-reviewed academic journals with high-impact factors wherein the proxy for ‘impact’ is the frequency with which other scholars cite a given work. Yet, this academic norm overlooks a simple truth:

Examination of research theories, methods, results, and policy proposals by contending interests in a real-world policy context can provide a more rigorous critical assessment of views than any provided in the more antiseptic atmosphere of professional journals and meetings. In the policy debate, the theoretical and methodological trappings of research are directly confronted by reality and the test of relevance (Melody and Mansell 1983, 113).

Put simply, there is a case to be made that: (i) the peer review process for OGP work actually is more thorough and – certainly – more transparent and accountable than is the case with much academic publishing; and (ii) that OGP/IRM research is as (and possibly more) impactful than the social science and humanities research whose relevance is oriented foremost to discipline-bound academic audiences.

In addition, OGP-related work often requires research activities to be conducted in a time-sensitive manner, resulting in highly accelerated data collection and writing timelines when compared to much traditional scholarship. Fast-paced work is vitally important, as it ensures IRM outputs have utility and align with OGP’s two-year program cycles. For example, reports detailing the strengths and limitations of a government’s public engagement activities in the co-creation of their open government action plans (i.e., Design Reports, and more recently, Action Plan Reports) need to go through both the rigorous multi-step internal IRM review and the external public review noted above. They also need to be made publicly available well in advance of the start of planning for the subsequent action plan so

that the issues raised can be addressed and lessons can be learned. In other words, the IRM's analyses and recommendations are meant to translate into action and change on the ground. To this end, governments are required to report back to the IRM in the form of self-assessment reports in which they address how they have incorporated feedback from IRM researchers into their next steps forward.

For those of us working in not-so-standard domains, the lack of recognition within the academy is doubly frustrating given the all-too-frequent disconnect between the supposed institutional priorities touted by university administrators – whether at the unit, faculty, and/or more senior administration levels – and how work that actually advances these priorities is weighted when it comes to career-related decisions. For instance, the University of Ottawa's strategic plan, *Transformation 2030: Building the University of Tomorrow*, sets out as one of its objectives, "Connect with local, national and global partners for research and knowledge mobilization" (University of Ottawa). Work such as the collaborative projects with OGP and the IRM clearly helps to build precisely these types of global connections. However, the objectives put forth in this plan – and many others like it from universities around the world – remain divorced from criteria for tenure, promotion, and other forms of recognition. There are two factors at play here. First are the challenges associated with operationalizing institutional priorities at the ground level. We suggest, without prejudice, that members of unit- and faculty-level academic personnel committees are not always aware of their university's five-to-ten-year strategic objectives and that such objectives seldom factor into the decision- or recommendation-making process. Second is the fact that decisions about promotion and tenure usually call upon the services of external academic experts who have little to no knowledge of the requesting university's strategic objectives. We recognize this gap is to be expected and, from an administrative perspective, may even be desirable. Nonetheless, the seeming inability – and, at times, what seems like a lack of motivation – to bridge this gap is an ongoing source of frustration for those whose work clearly is in line with, and advances the strategic priorities of, their institutions but perhaps less bound to adhere to disciplinary boundaries.

A third consideration regarding the disconnect between rhetoric and practice when it comes to formal recognition of this type of applied scholarship presumably stems from challenges associated with 'measuring' reach or impact. Historically, it has been more difficult to gather adequate metrics relating to work published outside the so-called hard sciences, let alone non-academic publications. Arguably, and despite known limitations with

citation counts, the h-index, and other related indicators, measurement is getting somewhat easier with the growing use of digital identifiers such as ORCID. Although the impact of non-typical academic work may be challenging to measure, researchers engaging in this type of scholarship know that such work has an impact in the real world, not least in bringing about changes in government policy and practice. Our OGP/IRM scholarship has directly contributed to augmenting the level and types of government-citizen engagement in decision-making across a variety of domains, spanning issues from the training of public servants to making it easier for Canadian citizens to access certain types of government data and information, to reforming beneficial ownership regulations. Given its reach, our efforts have garnered the attention of federal and provincial politicians in Canada, public servants, media, and the broader public. Our work has also resulted in a steady stream of invitations to speak to various groups within government and to give expert testimony to parliamentary committees. Additionally, there have been media interviews and offers to participate in conferences, workshops, and a host of other activities pertaining to matters of open data and open government.

Applied OGP-Focused Scholarship: The Professional Benefits

Invitations and attention to the work itself are only minor parts of what makes hands-on applied scholarship rewarding to researchers. In reflecting on our contributions to OGP, we have identified five other rewards that we detail below.

Front-Row Seats

Research and work, like that of OGP, requires frequent contact with practitioners and with the intended beneficiaries of the reforms being pursued and the changes being implemented. In our cases we have had to engage with public servants and members of civil society in Canada who are working in open government-related domains (e.g., Open Contracting Partnership, Transparency International, Open North). Our endeavors have also entailed working with OGP/IRM staff, other IEP members, other IRM researchers around the world, as well as government and civil society representatives from other countries. Such immersive and long-term efforts have provided us with front-row seats to what is going on in practice. From a researcher's perspective, the benefits of this longer-term, quasi-ethnographic immersion

in the workings of open government are immeasurable, not least because of the comprehensive and nuanced picture it provides of the intricate workings of the competing cultural, economic, political, social, and technological factors and tensions influencing action and inaction.

One illustrative example is the range of constraints that have complicated and, at times, impeded the design and implementation of national action plans and open government commitments in Canada. Three interrelated factors are particularly noteworthy here. First, responsibility for implementing and overseeing the country's OGP action plans falls under the remit of the Treasury Board Secretariat's (TBS) Open Government Team. However, it is individual federal departments and agencies – sometimes in collaboration with the TBS – that are responsible for delivering on action plan commitments and achieving milestones. Within this context, the Open Government Team's principal function is one of coordinator and facilitator, with much energy expended on getting department leaders and staff to understand how the latter fit into both their respective departmental briefs and the OGP process. Second, actions specified within commitments are designed and implemented in accord with the lead departments' and agencies' existing annual budgets and operating schedules. Most often, there are no departmental and/or agency budget lines specifically dedicated to delivering on OGP-related activities. As a result, thinking about action plans and commitment content tends to be oriented towards deliverables that fit in with existing departmental initiatives and budgets. Third, time is needed to secure the financing and the approvals required to implement ambitious open government initiatives. Successfully navigating the approval process takes, on average, some eight to eighteen months, and it is only after it has been successfully completed that a 'new' idea can be presented as a 'formal' commitment in an action plan. A lack of congruence between the time when 'new' ideas emerge, the government's budget cycle, and departmental operating schedules can, and often do, directly mitigate against the design and implementation of ambitious action plans and commitments. Put simply, given OGP's two-year program cycles, within the Canadian context the art of the possible is determined by government scheduling.

Misalignment between OGP program cycles and the timing of federal budgets makes it difficult for the federal government to associate open government commitments with budgetary commitments. As a result, open government activities set out in Canada's national action plans often end up as under-resourced projects that land on the desks of public servants, many of whom already have busy portfolios. This is precisely the type of nuanced insight that would escape us were it not for our front-row seats.

Network Building

Another advantage of having front-row seats is the network that gets built along the way. As previously mentioned, working with OGP has meant collaborating with a wide range of professionals from diverse sectors, both nationally and internationally. Building these connections helps to ground academic work. Indeed, so many of the academics, government, and civil society representatives involved with OGP found this to be so that, in response, they worked with OGP to establish 'Academic Days' that take place in conjunction with OGP's major international summits. These events are designed as typical academic conferences, complete with calls for papers, peer-reviewed paper selection, and, periodically, the publication of conference proceedings. There also is a pragmatic angle to these events. All too frequently, academics are unable to use their institutional and research funds to attend conferences at which they are not speaking. Participation in the Academic Days offers academics a means of circumventing such inane policies and to benefit from the network building and learning that coincides with attending OGP summits, even if they are not speaking or delivering a research talk at the summit itself.

The network, and the work itself, also contribute to fostering new lines of scholarly- and practice-oriented inquiry, insofar as hearing from other academics and practitioners helps to identify gaps in existing knowledge and emergent trends in open government practice. Contrary to some of the longstanding pejorative tropes outlined at the start of this chapter, nothing precludes one from researching, publishing, and advancing knowledge for audiences in both academic and practitioner domains.

Contributing to Methodological Rigor

Having opportunities to participate in discussions relating to research design and methodology is another rewarding facet of our work. Although there are foundational principles anchoring the understanding of open government, there is no one size fits all approach to implementing and evaluating open government initiatives. Hence, in establishing the eligibility of its members and in conducting the research associated with the IRM reports, OGP relies quite heavily on the use of indicators that have been developed by external parties. These include: the Open Budget Survey conducted by the International Budget Partnership, the Right to Information Rating (RTI) developed by Access Info Europe, the Democracy Index compiled by the Economist Intelligence Unit, and the World Bank's Officials

Financial Disclosure Database. For those engaged in this type of research, the combination of working with these and many other indices, and having the opportunity to investigate whether, and/or the extent to which, government claims of opening government stand up to empirical scrutiny contributes to developing of more in-depth and nuanced understandings of the links connecting theory, praxis, and the measurement of open government. It also helps to elucidate diverse and rich country contexts in which open government initiatives are being implemented.

OGP's IRM also has its own methodology for assessment that is used by IRM researchers and upon which they provide a steady source of feedback about the adequacy and reliability of the metrics used. This enables the large network of IRM researchers to generate reports that provide data that is comparable across countries. Here, the IEP plays an ongoing quality-control role, ensuring that the methodology is applied as consistently as possible. It also conducts regular reviews of the methodology with the aim of addressing issues flagged by IRM researchers. Working with the established methodology and with country-specific data and professionals allows IRM researchers to formulate unique insights into methodological limitations associated with monitoring and evaluating long-term, large-scale international governance-related, data-based initiatives.

One notable example of the limitations we have identified with the IRM metrics, and indeed much of OGP's general emphasis, is the propensity to focus only on the public-facing actions that governments undertake to advance open government. Changes in practice and/or work such as training initiatives that are inward-facing are not recognized as relevant to OGP's vision of open government. The result is a blind spot when it comes to recognizing that advances towards open government cannot be made without accompanying efforts in capacity-building. In other words, the current IRM methodology fails to recognize or otherwise account for the fact that it is public servants who ultimately are responsible for implementing and *doing* open government, and for engaging with external actors.

Creating Something Meaningful in Practice

One of the most rewarding outcomes of our work in this domain has been the opportunity to do, create, and write things that are meaningful in practice, with consequences that we witness with relative immediacy.

The IRM reports feed directly into the policy processes of OGP member governments. They are read and their recommendations are acted upon by governments and a wide variety of other civil society actors. The contents of

these documents also directly contribute to generating a range of datapoints that are curated by OGP in the form of an open-access database called *OGP Explorer*. This tool is particularly beneficial for anyone interested in conducting comparative and/or longitudinal studies of open government.

The IEP's work ensures quality and establishes the methodology for assessment of which member governments are highly aware. To this end, the creation, design, and implementation of national action plans tends to be developed with the assessment criteria in mind so as to maximize opportunities for favorable IRM reviews.

Creating Learning Opportunities for Students

The type of hands-on work we have been discussing creates myriad learning opportunities for students. For example, in 2016, Mary Francoli developed a new undergraduate course in open government at Carleton University. Since then, students enrolled in the course have had the opportunity to hear from a range of national and international guest speakers. They have been able to engage in their own applied research, working on projects in collaboration with community partners such as the federal Treasury Board Secretariat's Open Government Team, the Department of Public Safety, and the Canadian Security and Intelligence Agency, to name a few. They have also collaborated with OGP's communication team to write and publish blog posts relating to open government. In 2019, students from both Carleton and the University of Ottawa participated in the 6th Annual Global Summit on Open Government, held in Ottawa, and the coinciding Academic Day. Daniel J. Paré has likewise developed a graduate-level seminar in open government that is scheduled to be offered in 2025 to Master's and PhD students attending the University of Ottawa.

Outside the classroom, our work with OGP has contributed to research opportunities for both undergraduate and graduate students. These possibilities can take multiple forms. Students can be connected to networks built to help advance their own research projects. For example, at the time of writing, Daniel J. Paré is supervising a doctoral candidate whose dissertation research focuses on how transparency is being operationalized within the context of the Canadian government's public procurement activities and on whether its open contracting practices adhere to OGP principles and values. Students can be involved in the research and resulting publications of faculty research. Mary Francoli has also supervised several undergraduate and graduate theses directly related to open government.

She has hired a number of graduate students to support research projects on open government.

Such learning opportunities benefit students within the university setting and can help those seeking employment opportunities upon graduation. The skills garnered through applied research are highly marketable, as is specialist knowledge in the area of open government. Additionally, connections made with practitioners, either through guest speakers, assignments, or research opportunities, provide students with points of contact to draw upon when seeking employment.

Conclusions

It is clear that applied scholarship, like the hands-on OGP-related work detailed in this chapter, feeds and enriches all elements of academia, including research, teaching, and service. Unfortunately, institutional norms, traditions, and practices continue to mitigate against fully recognizing the value and merit of such work. This said, concrete measures can be taken at the institutional and individual levels to help in this regard.

At the institutional level, revisiting the criteria applied in the making of career-impacting decisions such as tenure and promotion, and recognizing the merits of different types of scholarship, is important. Carleton University, home to one of the chapter's authors, took a step forward in this regard around 2012 when it introduced more generic language about general university levels of achievement. In so doing, it created a means of moving forward with myriad interpretations of achievement and differing unit-level standards for tenure and promotion. At present, these include unit-specific indicators that are drawn up and agreed upon by individual units, external peer reviewers, the Faculty Dean, and the Provost. This model provides units an opportunity to build in consideration of a wide range of work, including innovative teaching that connects students with the community, research funding obtained from sources other than Canadian Tri-Council grants, policy work, and professional reports such as IRM reports. These standards are provided to all committees within the university, and to external reviewers considering tenure and promotion applications, helping them to read and adjudicate applications. They also provide the applicant with clear guidance on institutional expectations.

Mentorship of junior faculty, in particular, can be instrumental in determining how to best make and communicate linkages between traditional and applied scholarship. This is vital when communicating one's work inside

and outside the academy, not least because the way one speaks about one's work necessarily differs according to the audience being addressed.

At the individual level, scholars engaging in the type of applied work described in this chapter need to think carefully about the narratives they use to describe their work. This is particularly vital in tenure and promotion applications and when applying for new positions. Institutional mentorship and support, such as workshops and one-on-one guidance, can help to ensure that applications are strong. Individuals also have a growing toolkit of technology that can be used to identify their work and to bolster its likelihood of recognition. Digital identifiers, such as the ORCID which applies to a wide range of work outside the scope of traditional refereed journal articles, manuscripts, and book chapters, can help researchers in tagging and receiving credit for their work.

In closing, hands-on applied research is often a labor of love and it can be immensely rewarding. It provides researchers with front-row seats to what is happening in practice, it builds invaluable networks, it contributes to methodological rigor, it leads to the creation of something that is meaningful, and it can provide enriching learning opportunities for students. Nevertheless, it continues to be undervalued within academia. There are, however, signs that this is changing, and there clearly are institutional and individual steps that can be taken to help in this regard. As with the implementation of open government, successfully bringing about this change on a wider scale will, of course, be contingent upon the political will to transform organizational culture.

References

- Borden, Kyla. 2017. *Raising the Open Government Bar: Changes to Canada's Access to Information Act*. Retrieved from: <https://www.opengovpartnership.org/stories/raising-the-open-government-bar-changes-to-canadas-access-to-information-act/>.
- Carey, James. 1978. "The Ambiguity of Policy Research." *Journal of Communication* 28 (2): 114–119.
- El-Sammak, Katherine. 2016. "Measuring Open Government: Critically Reviewing OGP's Footprint." November 17, 2016. <https://www.opengovpartnership.org/stories/measuring-open-government-critically-reviewing-ogps-footprint/>.
- Francoli, Mary. 2015. *Canada Progress Report: 2012–13*. https://www.opengovpartnership.org/wp-content/uploads/2019/07/Canada_final_2012_Eng.pdf.
- Francoli, Mary. 2016. *Canada Progress Report: 2014–15*. https://www.opengovpartnership.org/wp-content/uploads/2017/05/1.Canada14-15_English_Final_o_o.pdf.

- Francoli, Mary. 2017. *Canada End of Term Report 2014–16*. https://www.opengovpartnership.org/wp-content/uploads/2017/05/Canada_EOTR_2014-2016_EN.pdf.
- Fraundorfer, M. 2017. "The Open Government Partnership: Mere Smokescreen or New Paradigm?" *Globalizations* 14 (4): 611–626. <https://doi.org/10.1080/14747731.2016.1236463>.
- Hamilton, Malika. 2019. "Engaged Youth: Let's Talk." March 25, 2019. <https://www.opengovpartnership.org/stories/carleton-university-student-blog-series-engaged-youth-lets-talk/>.
- Jaeger, Paul T., and John Carlo Bertot. 2010. "Transparency and Technological Change: Ensuring Equal and Sustained Public Access to Government Information." *Government Information Quarterly* 27 (4): 371–376. <https://doi.org/10.1016/j.giq.2010.05.003>.
- Jensenius, Francesca, Mala Htun, David J. Samuals, David A. Singer, Adria Lawrence, and Michael Chwe. 2018. "The Benefits and Pitfalls of Google Scholar." *American Political Science Association* 51 (4): 820–824. <https://doi.org/10.1017/S104909651800094X>.
- Lazarsfeld, Paul, F. 1975. "The Policy Science Movement (An Outsider's View)." *Policy Sciences* 6 (3): 211–222.
- Melody, William, H. and Mansell, Robin, E. 1983. "The Debate over Critical vs. Administrative Research: Circularity or Challenge." *Journal of Communication*, 33(3): 103–116.
- Paré, Daniel. 2018. *Ontario Progress Report 2017*. https://www.opengovpartnership.org/wp-content/uploads/2018/10/Ontario_Final-Report_2017.pdf.
- Paré, Daniel. 2021. *Canada Design Report 2018–2020*. https://www.opengovpartnership.org/wp-content/uploads/2021/07/Canada_Design_Report_2018-2020_EN.pdf
- Pastran, Camila. 2016. *Open Government: A Solid Step Towards Peace Agreements*. <https://www.opengovpartnership.org/stories/open-government-a-solid-step-towards-peace-agreements/>.
- Reiter, Dan. 2016. "Citation Count Data and Faculty Promotion." *Duck of Minerva*. February 18, 2016. <https://www.duckofminerva.com/2016/02/citation-count-data-and-faculty-promotion.html>.
- Stephan, Paula, Reinilde Veugelers, and Jian Wang. 2017. "Reviewers Are Blinkered by Bibliometrics." *Nature* 544 (7651): 411–412.
- University of Ottawa. *Transformation 2030: Building the University of Tomorrow*. https://transformation2030.uottawa.ca/sites/default/files/DGC19_1135_T2030_BROCHURE_ENG_WEB4.pdf.
- The White House. 2010. "Remarks by the President to the United Nations General Assembly." September 23, 2010. <https://obamawhitehouse.archives.gov/the-press-office/2010/09/23/remarks-president-united-nations-general-assembly>.

About the Authors

Mary Francoli is the Associate Dean and Director of the Arthur Kroeger College of Public Affairs and Associate Professor of Communication and Media Studies at Carleton University. She has been a long-standing member of the Open Government Partnership's International Expert Panel and is co-chair of Canada's National Security Transparency Advisory Group.

Daniel J. Paré is an Associate Professor of Communication and Media Studies with the University of Ottawa's Department of Communication and holds cross-appoints with the university's School of Information Studies (ÉSIS), and its Institute for Science, Society and Policy (ISSP). His research examines governance challenges arising from innovations in information and communication technologies.

6. The Challenge of Addressing Subjectivities through Participatory Action Research on Datafication

Katherine Reilly and Maria Julia Morales

Abstract

Participatory action research leans heavily on the subjectivity of knowledge producers as a vehicle for change. In our work, however, we encountered several assumptions about how subjectivities, including our own, can be drawn on to address datafication's implications for data subjects. We argue that participatory data literacy interventions, if they are to create fundamental change in the face of datafication, cannot be technocentric or top-down instruments. Rather, they must depart from the processes of subjectivation taking place in relation to specific instances of datafication. In addition, data subjects must have the tools necessary to engage in critical reflection so that they recognize collective subjectivities and/or embrace their individual agentic capacity. Without these tools, it is difficult to realize participatory transformations.

Keywords: Data subjects; Subjectivity; Citizen data audits; Data literacy

Introduction

From September 2019 to December 2022, we worked with partners in five Latin American countries to realize participatory data literacy interventions with people who were sharing their personal data during daily activities such as shopping, banking, or seeing the doctor. These people are commonly referred to as 'data subjects,' meaning that they can be identified through the data they share. However, certain critical perspectives call

this conception into question, pointing out that our subjective experiences of existing, expressing and exerting ourselves are social constructs, the products of historical and structural forces including complex experiences with datafication.

These observations complicate the idea of participatory action research, which leans heavily on the subjectivity of knowledge producers as a vehicle for change. By 'subjectivity' we mean people's cognitive and emotional positionality within our heavily mediated sensorium (building on Martín-Barbero and Rincón 2019); 'subjectivization' refers to the processes of social construction and consciousness-building that happen within such a space. In this sense, we understand a sensorium to be a regime of perception. In our work we encountered several assumptions about how subjectivities, including our own, can be drawn on to address datafication's implications for data subjects. Thus, in this chapter we draw on our experiences to problematize the application of participatory action research to datafication.

Our project brought together academic investigators, civil society partners, and participants drawn from the public. It was organized into three interrelated processes that advanced roughly in parallel. Our partner organizations, which included Derechos Digitales from Chile, Hiperderecho from Peru, Karisma from Colombia, Tedic from Paraguay, and ObservaTIC from Uruguay, each identified a local case study and mobilized people affected by a particular issue there to participate in a data literacy intervention. Meanwhile, the research team carried out an intensive literature review focused on theories of datafication, information systems, mediation, and platformization, drawing especially on work by scholars from Latin America. We then used these perspectives as inspiration for a set of techniques that grounded learning in participants' daily experiences, personal values, needs, desires, and perspectives, as well as collective processes of reflection. The decision to focus on everyday experiences with datafication was a response to a 2018 pilot project which found that, because datafication and subjectivization are relational and situated phenomena, universal data literacy strategies are unlikely to create fundamental change (Reilly 2020).

The intersection of our various roles while carrying out this work revealed important assumptions about the nature of subjectivity under datafication. We struggled to engage in dialogue about our project goals given the transnational, cross-cultural, and multisectoral nature of our work. What did we each bring to the work, and how could we surface this and negotiate differences as a foundation for participatory action research? We needed to negotiate the demands of the various actors involved, but it was challenging to do so given our very different perspectives on datafication

and its relationship to subjectivity. We worked alongside our partners and participants to establish methodologies that balanced local needs with project goals and introduced ways to engage with users' specific experiences and views. This intensive work was important because when it comes to data literacy, people's daily experiences have little to do with data itself but rather rest on the creation of meaning that arises out of datafied transactions, which both emerges from and results in subjectivization.

Through this work, we've come to identify two areas for further engagement. First, data literacy interventions, if they are to create fundamental change in the face of datafication, cannot be technocentric or top-down instruments but rather must depart from the processes of subjectivization taking place in relation to specific instances of datafication or platformization. Second, data subjects must have the tools necessary to engage in critical reflection so that they recognize collective experiences of subjectivization and/or embrace their individual agentic potential. Without these tools, it is difficult to image participatory transformations through either cultural change or advocacy channels.

Data Subjects and Participatory Action Research

Participatory action research, as traditionally understood (Fals Borda 1987), is a relational methodology that leverages ties created among the different actors involved in a knowledge production process to create change in the world. Researchers, practitioners, and community members might come together to produce knowledge that supports policy advocacy, for example. By this logic, we either assume that actors' subjectivities are fixed and advocacy work benefits from the sum of their contributions to change, or that subjectivities evolve through these interactions in ways that both encompass change and also create it. In some perspectives, these collaborations involve the embrace of critical consciousness as a means to enable change. In either case, the knowledge that results from action research is not an objective truth but rather a historically and culturally situated construct that emerges in the context of a particular goal, given the relationships that evolve among the participants and the processes through which this knowledge is produced. In turn, this knowledge changes the society it refers to, either directly, by informing decision-making, or indirectly, by transforming the knowledge sphere and the relations within it.

This cohesive image of action research is challenged by Dubet's (1996) idea of the 'neomodern' subject. Dubet argues that in our contemporary world,

the unity of our roles has been displaced by the diversity of our experiences, and that our subjectivity emerges in and through the administration of the various logics we encounter in the many fractured expressions of modern life. Rather than manifesting cohesive subjectivity, he argues, it is more accurate to say that in our contemporary world, we search for authenticity and engage with meaning-making as we navigate and negotiate the many autonomous systems that we encounter on a daily basis.

This notion is reflected in critical and engaged research about information systems and personal data. For instance, Martín-Barbero (2011) explored how, in a digitally networked era, “the body itself is inserted into life worlds and into the instability of senses that configure an individuality that suffers from constant identitary fragility” (116). He argued that with the collapse of church and state as regulators of culture, “both individual and collective identities are subject to the fluid oscillation of guides and their interpretation, adjusting themselves to the image of a fragile network, without center and in continual movement” (ibid.). Similarly, Sefton-Green and Pangrazio (2021) argue that we are confronting the ‘death of the subject’ given the ‘regimes of subjectivization’ created by myriad processes of datafication. They argue that “datafication proliferates the divisions of the singular, autonomous person into the ‘dividual’” (7). The ‘dividual,’ they explain, is a concept put forward by Smith (2012), which captures the idea that in our multiply mediated world, our experience is made up of various separable but interdependent dimensions. This situation culminates in the creation of a fractured individual who lacks a clear sense of self and the necessary foundations for critically engaged citizenship.

To the extent that this is true, both Dubet and Martín-Barbero argue that we can no longer locate the idea of society in social contracts or ideological domination. Rather, society is to be found in diverse and fractured processes of meaning-making mediated by the various communications technologies that create the contemporary conditions for our lifeworld. This would suggest that the lifeworld is at risk of colonization given the power of information systems and datafication to influence meaning-making, an observation which is central to action research about datafication. The problem is that we cannot rely on clearly expressed subjects to be a starting point for actions to create change, because those same systems of meaning-making exert their influence through fractured and shifting immediacies.

These observations raise important questions and concerns for critically engaged research that seeks to create change in the world. For starters, we need to query the extent to which data research is, itself, a force for subjectivization. Are our data-centric action research projects creating, or

are they reifying, information systems that contribute to our participants' fractured experiences? And should action work in the field of data studies seek, as its goal, to work with data subjects so that they can recuperate subjectivity, embrace agency, and reclaim their sense of self (Touraine 1994)? Much data literacy work operates in precisely this vein (Markham 2019; Pangrazio and Selwyn 2019). However, this type of work is challenged by postcolonial critiques which suggest that research should not, indeed cannot, intervene to construct subjectivities, because subaltern consciousness emerges autonomously from outside the logics of the system; research is appropriate only to the extent that it engages in rearguard actions (Santos 2015) that serve the needs and desires of affected communities. And yet it is not at all clear that a 'subaltern consciousness' can still be said to exist beyond the boundaries of contemporary systems of mediatization and datafication. If such consciousness does not in fact exist, then there is arguably a need for work that tracks the implications of information systems and datafication for identity formation and meaning-making, as well as reflection about appropriate methods, processes, and forms of collaboration.

Having said all this, Livingstone (2019) argues that we need not be so concerned: the reach of information systems is not as hegemonic as popularly claimed by observers of datafication. As she says, "not all public connection is significantly mediated, and nor is all mediated experience determining of democratic participation" (178). In this assessment, there is still room for audiences and publics to interpret their reality and to produce autonomous assessments of their experiences with datafication. Similarly, Martín-Barbero was convinced of communities' power to resist and reinterpret the technical force of digital information systems (in Couldry 2019, 187). Drawing on Martín-Barbero's work, Rincón (2019) argues that affect theory offers a means to "read the analytics and comprehension of contemporary modes of perception and feelings created by technicities" (270). For example, surveillance technologies can be understood based on how they make people feel. In this view, community members express their expertise about datafication through the fear, mistrust, frustration, vulnerability, insecurity, desire, agency, creativity, and resilience that they experience in their everyday lives (Lewis et al. 2018). This, then, forms the basis of critical consciousness, creative engagements, or a resistance to invasive, profiling, and discriminatory technologies.

A focus on people's lived and affective experiences allows participatory action researchers to sidestep the dilemma of subjectivizing data subjects. However, further challenges present themselves in the analysis of affective expressions. Each set of experiences emerges in the context of a single

individual's interactions with a single information system. To complicate matters, the content of these emotions and experiences is situated in relation to a variety of personal, organizational, or national relationalities. As Sefton-Green and Pangrazio (2021) point out, "the version of the self on Instagram needs to be different to the one that appears on their school's learning management system in order to fit in and/or comply with different standards and benchmarks" (7). From the point of view of participatory action research, this means that researchers are no longer in dialogue with subjects who occupy roles; we are no longer bringing together everyone's situated expertise to collaboratively solve a problem. Rather, we are more like therapists or pedagogues engaged in dialogue with desires and experiences that may or may not align with individual and group subjectivities. As a result, it isn't clear whether action research about datafication should work to acknowledge and enable subjective experiences within a fractured mediated plain (Rincón 2018), transcend the 'dividual' in an effort to reintegrate subjectivities (Sefton-Green and Pangrazio 2021), or bring individual experience into dialogue through care work (Skipper and Pepler 2020) as a basis for community-building.

Action research with data subjects has confronted these challenges in different ways. Early efforts simply revealed the volume and character of data collected to users of information systems (Lupton and Michael 2017; Selwyn and Pangazio 2018). Investigators were surprised and dismayed to discover that their revelations carried little weight with participants. Rather than expressing shock at breaches to the integrity of their being, they calmly observed the banality of data flows. Subsequent efforts sought to engage individuals in meaning-making through critical reflection about datafied experiences. For example, Markham (2019), in her work on critical pedagogies of datafication, argues that research findings "are not for me, as the facilitator, but for the users who become researchers themselves. What they do with these findings is up to them" (756). She continues: "The research outcome for me is not to provide answers but to raise questions and cause a chain reaction whereby participants raise their own questions and ask their parents, siblings, friends, and colleagues to also raise questions" (ibid.). Yet, in a more recent project, Markham (2021) grapples with the challenge of discursive closure around datafication and how it creates a sense of inevitability that prevents data subjects from imagining alternative futures. In both of these studies, she focuses on facilitating individual self-reflection and self-awareness, but people are not mobilized for any sort of communal or external change. In contrast, novel theorizations of action research seek to move beyond critical consciousness and engage in processes that generate

empathy and respect (Skipper and Pepler 2020). In this conception of action research, the goal could be to build bridges between various experiences of datafication with the aim of generating communities of understanding.

At the same time, a growing push to “reduce the burden of digital literacy” on individual citizens (Livingstone 2019, 178) has emerged, to be achieved through regulation that constrains the extremes of datafication and mediatization. By making the balance of power between digital platforms and data subjects more equitable, it is hoped that people will have better opportunities to express their citizenship (Kennedy and Moss 2015). Of course, such policies would make the playing field more fair, but there is a catch-22. If people lack digital literacy and data consciousness, how can they participate in policy-making processes in ways that reflect their local experiences with datafication, as well as their values, needs, and desires? And even if they do possess a greater ability to monitor data use, will they actually do so if the information systems they are involved in contribute to a fracturing of their subjectivity and sense of belonging to a larger social project?

The question then becomes how to support self-reflection that helps people engage more deliberately in the processes of meaning-making that shape their own particular social contexts, and perhaps also become conscious of the need to change the conditions under which they carry out these acts. This is what we endeavored to do through our data literacy interventions.

Datafication and Subjectivity as Seen Through Our Work

The landscape for action research around datafication described above served as the backdrop for our data literacy interventions in Latin America. As we addressed the challenges we faced in the field, we came to better understand the role of subjectivities in participatory action research about datafication.

Our first dilemma arose with the identification of participants for our study. As principal investigators, we worked with our partner organizations to identify information systems whose data practices have important implications for the wellbeing of users. Once this was established, our partners needed to locate participants from these user groups to take part in the study. In four out of five cases, the groups that emerged were not natural communities sharing a common subjectivity but rather were made up of diverse individuals, each of whom happened to be related to or affected by

the information system in question. For example, Observatic brought together senior citizens who must use the same information system to access medical care but whose subject positions vary greatly in terms of socioeconomic status, education, and political commitments. However, Karisma decided to work with a union for a gig economy company called Rappi, so here there was a clearly defined subject with strong potential for action. In some cases, partners worked on the informational practices of a sector rather than a single information system. For example, HiperDerecho worked with women who purchase contraceptives from pharmacies, which led to discussions about all sorts of information systems, ranging from the informal knowledge exchanges resulting from one-on-one conversations to the loyalty programs that provide discounts to regular shoppers at these stores.

What appeared at first to be a failure of participatory action, or a roadblock for research, ended up revealing the diverse ways people experience information systems. Groups did in fact emerge through our processes, and participants discovered commonalities in their values, experiences, and desires, but they have not necessarily located a common subjectivity or basis for collective action. What is more, each project brought something different to the fore – gender, migration, the elderly, youth, and workers – each within a different local context. Context, it turns out, matters deeply for people's experience of datafication. Take the respective cases of Uruguay and Peru, for example. In the former instance, people's suspicions and fears about data and privacy are intimately related to the repression and surveillance that occurred during the country's authoritarian military regime. In the latter, datafication has emerged in the context of the neoliberal economic stabilization that followed a protracted civil conflict, so inclusion in information systems is associated with status and security. Hence there is no single or clear 'data subjectivity' for action researchers to work with. It is not entirely certain whether this matters in terms of mobilizing change, but as action researchers we were left with questions about how change would happen.

Indeed, as the project continued, it became evident that researchers and partners had very different views of how our work would play out. As researchers, we initially imagined our work to be engaged in consciousness-raising, enabling participants to contest how datafication was shaping their perspectives and experiences. When a participant from Uruguay chidingly asked, "So what do you do with the data that we've shared with you?", we celebrated: datafication was being questioned! Similarly, participants in Paraguay were impressed by our informed consent letter and used it as an example of how corporate actors should work with personal data. We saw this as evidence that our participants were becoming more critical about

data, could better identify what is at play in a data transaction, and were better able to express their own values in personal decision-making. And yet we are left with an unsettled feeling that the project did not achieve the impact we had hoped for at the level of information systems or policy change.

Meanwhile, our partners are digital rights organizations who actively participate in international discussions about data justice and work nationally to effect change in the legal context of information access, technology use, privacy, copyright, net neutrality, and similar concerns. As people on the front lines of digital policy issues, they approached communities as experts. They wanted to educate people about their rights and mobilize these people's agency as citizens to exercise these rights, with little consideration for either subjectivity or experiences.

In addition, the project initially imagined that it would develop technological interventions to support the work of its participants. Specifically, we planned to identify people's situated values *vis-à-vis* data use and information systems, as well as their main preoccupations and concerns about them. We then planned to use these identified values as a basis for the design and development of digital tools that would help groups conduct audits of the data use of corporate information systems. These plans suited our partner organizations, because as NGOs they liked the idea of creating and distributing a tool that citizens could use to exercise their rights. However, as the project took shape, it became clear that what citizens lack is not the ability to exercise their rights but rather the time, space, and suitable process to reflect on their experiences. We found ourselves developing workshop techniques for group activities rather than technological interventions in support of audits.

These differences in expectations and shifts in project goals meant that we had to negotiate demands with our research partners, which involved surfacing different ways of thinking and the making of compromises. One particular site of discussion emerged around definitions of data. Our partners wanted to begin their community engagement by sharing the legal definition of data rather than by exploring how participants experienced and understood data. Partners also had a tendency to focus on the value of data as a commodity rather than on people's values *vis-à-vis* data.

Eventually we decided to hold a workshop to highlight our different ways of thinking about data: namely, as the 'grease' that facilitates transactions and that therefore has an exchange-value, as opposed to a socially and culturally constructed artefact that reflects ethical commitments, traditions, community needs, and political interests. We explored examples from Indigenous data sovereignty movements (Animiiki 2019), African American

data rights work (McNealy 2020), and data feminism (D'Ignazio and Klein 2020) to illustrate the latter. We concluded that participatory interventions should surface and explore participants' values, perhaps including but not limited to transactional understandings of data. We needed to start by seeing where participants were coming from. This would allow us to understand how participants conceived of data, how they express their subjectivity *vis-à-vis* information systems, and how that subjectivity informed individual and collective experiences.

This episode revealed important dimensions of NGO-led data work in Latin America that bear greater scrutiny. Our partners are professionalized organizations that move in transnational circuits of dialogue about data rights. The spaces they move in, such as RightsCon and the Internet Governance Forum (IGF), tend to emphasize legalistic and technocentric frameworks for thinking about data, and they also tend to prioritize the transactional nature of data. As a result, data rights work tends to position people's primary subjectivity as that of citizens and consumers; datafication's primary effect on individuals, it is presumed, is to undermine democratic participation and expose individuals to abuse in the marketplace. But starting from this set of assumptions reifies – at best – class relations by presuming that all people are equally included in national projects as citizens and consumers, which is manifestly not the case. At worst, these presuppositions push aside any careful examination of how subjectivity is actually expressed and gloss over the difficult questions about action research and subalternity raised in the previous section.

As we struggled with these challenges, we read Latin American sociological and communications literature to make sense of what we were observing and experiencing. We came to draw on Martín-Barbero's technical-cultural construct of agency/subjectivity. In this view, 'technicities' describe how information systems structure processes of meaning-making, knowledge production, and, especially, how we know and express ourselves. As he said, "If during centuries technology was considered a mere instrument, today it is already well on its way to becoming reason [itself]" (2011, 110). Technicities are balanced by 'sensorialities' as the space of sensations through which we experience our surroundings. As Martín-Barbero and Rincón explain, "we need the human, cultural and emotional to counteract emotionless informational apparatuses" (2019, 20). Rincón introduces affect into this formulation as "the force and potential for developing the wisdom of an epoch given its technicities", recognizing that "for their part, technicities make present new dimensions of sensoriality" that shape our affective experience (2019).

Thus, the main products of our work are techniques that people can use to make sense of their situated experience of information systems and datafication. For example, *¿Qué es un nombre?* (What is a name) asks participants to list different names they go by, the situations in which these names are used, how these name-situations make them feel, and the values they attach to them. For example, a family nickname might be used only by siblings within family settings and could make a person feel both nostalgia and paternalism. These prompts are designed to show the relationship linking naming practices, information systems, identity, and social relations. Auditors are then asked to reflect on how a particular information system identifies them, and what this identification suggests for their relationship with that system or other people within it. Further prompts explore how the ritual of name-sharing organizes social relations in the community. The goal here is to spark reflection about complex responses to a specific moment of data collection and about how to exit or negotiate the expectations of conformity embedded within the designs of those systems.

Similarly, *El Desborde* (The Overflow) is inspired by the work of Villasante (1999), who studied how identities are expressed in response to the many possible demands we face in daily life. This method asks participants to describe a specific instance of data sharing and reflect on what was expected of them in that moment, along with how they felt, how they reacted, and how they *wanted* to react. Auditors then contrast their actual participation in the information system with how their identity, needs, desires, or emotions ‘overflowed’ the requirements of that system. Villasante argued that people might respond to demands in various ways: they might comply with the system despite their misgivings; reject the system outright and forgo its services; engage in subterfuge by, for example, submitting false data; or identify gaps between their needs and what the system provides, and then demand change. Of these possibilities, he advocated the fourth option, which he saw as the best opportunity to redress injustices and improve the ability of systems to meet the needs of communities.

These tools offer data subjects ways to construct social meaning, express public values, and make sense of their current realities. However, we feel concerned that they might reify the very regimes of subjectivization that need to be superseded if people are to go beyond merely adjusting their individual practices. Sometimes people have no option but to make use of information systems, and it worries us that they might become more aware of how information systems are positioning or exploiting them without encountering some means to respond. We are left questioning how subjectivities can be reclaimed given the force of datafication’s technicities.

How do we move past individual experiences of datafication and reclaim subjectivization as a communal process of identity formation? How do we reduce the burden on data literacy as a means to protect individuals within digital spheres, without taking away people's agency as actors demanding new forms of mediation that respect local cultures, values, and desires?

Conclusions

Subjectivities emerge through the very processes of participatory action research that brought us together – as researchers, digital rights organizations, and participants – to produce knowledge. As Dubet (1995) argued, the subject must be defended and reclaimed through the articulation of experiences because, in contemporary society, this is what gives coherence to identity, strategies, and subjectivity. Similarly, Touraine wrote:

The individual only becomes a subject by separating from themselves, opposing the logic of social domination in the name of a logic of freedom, the logic of free production of one's self. It is about rejecting an artificial image of social life seen as a machine or organism. [...] in a world where humans have been transformed into objects, the free production of oneself affirms the subject and their rights. (Touraine 1992, 231)

As our participants and partners came together and engaged in dialogue with one another, not only did they identify their subjectivities, they also became actors with the potential to transform their world. For Touraine (1992) actors do not just act according to a role but rather can transform their reality's material structures – including the division of labor, the criteria for decision-making, and the relations of power or cultural orientations – through the expression of that role.

Our engagements have also forced us to rethink our research practices for the digital age. In particular, we learned the importance of setting aside technocentric sensibilities and of shifting data from being an object of study to a context for analysis. Social processes – ways of thinking, ways of doing, human relations – are what drive the transactions between the 'users' whom platforms, information systems, and software involve. This type of work is not straightforward. Analysis and intervention must be contextualized, which means that there cannot be efficiencies of scale. Adequate funding is required to support these types of engagements.

Researching these processes raises questions about how to engage with the subjectivities that come into play. We need new methods and tools for this type of research, particularly those that simultaneously engage people's experiences of datafication while also developing digital literacies. The former offer a means to produce concrete research outputs from community engagements, while the latter builds agency within the community itself. In this way, people can give meaning to the actions and strategies they use in the digital world, with a view to effecting transformations that better reflect their values, desires, and needs.

Acknowledgements

This work was supported by the Social Sciences and Humanities Research Council (SSRHC) and the International Development Research Center (IDRC), and it benefitted from the contributions of our research team, which includes Carolina Carretero, Ramiro Liesegang, and Esteban Morales.

References

- Animiiki. 2019. "Decolonizing Digital: Contextualizing Indigenous Data Sovereignty." <https://animikii.com/news/decolonizing-digital-contextualizing-indigenous-data-sovereignty>
- Couldry, Nick. 2019. "Afterword." In *Citizen Media and Practice: Currents, Connections, Challenges*, edited by Hilde Stephansen and Emiliano Treré. London: Routledge. 181–189.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Dubet, Francois. 1996. "¿Ocaso de la idea de sociedad?" *Revista de Sociología* (10): 7–23. <https://doi.org/10.5354/0719-529X.1996.27653>.
- Fals Borda, Orlando. 1987. "The Application of Participatory Action-Research in Latin America." *International Sociology* 2 (4): 329–347.
- Kennedy, Heather and Giles Moss. 2015. "Known or Knowing Publics? Social Media Data Mining and the Question of Public Agency." *Big Data & Society* 2 (2). <https://doi.org/10.1177/2053951715611145>.
- Lewis, Tamika, Seeta Peña Gangadharan, Mariella Saba, and Tawana Petty. 2018. *Digital Defense Playbook: Community Power Tools for Reclaiming Data*.

- Our Data Bodies. http://www.odbproject.org/wp-content/uploads/2019/03/ODB_DDP_HighRes_Single.pdf.
- Livingstone, Sonia. 2019. "Audiences in an Age of Datafication: Critical Questions for Media Research." *Television & New Media* 20 (2): 170–183.
- Lupton, Deborah, and Mike Michael. 2017. "Depends on Who's Got the Data': Public Understandings of Personal Digital Dataveillance." *Surveillance and Society* 15 (2): 254–268.
- Markham, Annette. 2019. "Critical Pedagogy as a Response to Datafication." *Qualitative Inquiry* 25 (8): 754–760.
- Markham, Annette. 2021. "The Limits of the Imaginary: Challenges to Intervening in Future Speculations of Memory, Data, and Algorithms." *New Media and Sociology* 23 (2): 382–405.
- Martín-Barbero, Jesus. 2011. "La pertenencia en el horizonte de las nuevas tecnologías y de la sociedad de la comunicación." In *Sentido de pertenencia en sociedades fragmentadas: América Latina desde una perspectiva global* by Ana Sojo, and Martín Hopenhayn, 105–127. Buenos Aires: Siglo Veintiuno.
- Martín-Barbero, Jesus, and Omar Rincón. 2019. "Mapa insomne 2017: Ensayos sobre el sensorium contemporáneo. Un mapa para investigar la mutación cultural." In *Un nuevo mapa para investigar la mutación cultural: Diálogo con la propuesta de Jesús Martín-Barbero*. Quito: CIESPAL.
- McNealy, Jasmine. 2020. "An Ecological Approach to Data Governance." Data & Society Research Institute. https://youtu.be/jB5_NrdWH7k.
- Mejias, Ulises A., and Nick Couldry. 2019. "Datafication." *Internet Policy Review* 8 (4). <https://doi.org/10.14763/2019.4.1428>.
- Pangrazio, Luci and Neil Selwyn. 2019. "Personal Data Literacies': A Critical Literacies Approach to Enhancing Understandings of Personal Digital Data." *New Media & Society* 21 (2): 419–437. <https://doi.org/10.1177/1461444818799523>.
- Reilly, Katherine. 2020. "The Challenge of Decolonizing Big Data through Citizen Data Audits." *Datactive*. <https://data-activism.net/2020/04/bigdatasur-the-challenge-of-decolonizing-big-data-through-citizen-data-audits-1-3/>.
- Rincón, Omar. 2018. "Mutaciones bastardas de la comunicación." *MATRIZES* 12 (1): 65–78.
- Rincón, Omar. 2019. "Mi invención: Sobre el mapa para comprender el sensorium de la contemporaneidad." In *Un nuevo mapa para investigar la mutación cultural: Diálogo con la propuesta de Jesús Martín-Barbero*. Quito: CIESPAL.
- Santos, Boaventura de Souza. 2015. *Epistemologies of the South: Justice against epistemicide*. London: Routledge.
- Sefton-Green, Julian and Luci Pangrazio. 2021. "The Death of the Educative Subject? The Limits of Criticality Under Datafication." *Educational Philosophy and Theory*. <https://doi.org/10.1080/00131857.2021.1978072>.

- Selwyn, Neil, and Luci Pangrazio. 2018. "Doing Data Differently? Developing Personal Data Tactics and Strategies amongst Young Mobile Media Users." *Big Data & Society* 5 (1): 2053951718765021. <https://doi.org/10.1177/2053951718765021>.
- Skipper, Yvonne, and Debra Pepler. 2021. "Knowledge Mobilization: Stepping into Interdependent and Relational Space Using Co-Creation." *Action Research* 19 (3): 588–605. <https://doi.org/10.1177/1476750320960810>.
- Smith, Karl. 2012. "From Dividual and Individual Selves to Porous Subjects." *The Australian Journal of Anthropology* 23 (1): 50–64.
- Touraine, Alain. 1994. *Crítica de la modernidad*. Buenos Aires: Fondo de Cultura Económica.
- Van Dijck, José. 2014. "Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology." *Surveillance & Society* 12 (2): 197–208.
- Villasante, Tomás. 1999. "Cuatro redes para hacer transformaciones sustentables." *Política y Sociedad* 31: 37–54.

About the Authors

Katherine Reilly is an Associate Professor in the School of Communication at Simon Fraser University. Her research explores participatory and community engaged methods to bridge the gap between human or community observations and the representations that get carried into decision making processes.

Maria Julia Morales holds a Doctorate in Educational Technology from the Rovira y Virgili University and MA from the Open University of Catalonia. She is a researcher with the ICT Observatory (ObservaTIC) at the University of the Republic, Uruguay. Her work explores appropriation of digital technologies and digital competencies.

Part II

Case Studies

7. Community Responses to Family Violence Policy: A Public Sector Collaboration

Anthony McCosker, Jane Farmer, and Arezou Soltani Panah

Abstract

In 2015–2016, the state government of Victoria, Australia, mounted a high-profile public inquiry into family violence, seeking input from various stakeholders with the aim of guiding policy reform and widespread systemic and social change. In subsequent years, the government sought to use innovative data science methods to evaluate the impact and outcomes of the inquiry and of family violence policy interventions. This chapter reports on a collaboration between an interdisciplinary team of social and data science researchers and public sector stakeholders using innovative digital methods and natural language processing techniques to provide evidence of public responses to family violence policy. It explains the opportunities, processes, and challenges involved in a collaborative public sector social data project.

Keywords: Social media data; Domestic violence; Natural language processing; Policy

Introduction

This case study tells the story of a research collaboration with an Australian state government that led a large-scale Royal Commission inquiry and enacted significant policy changes to combat family and domestic violence.

The aim of the work was to use social media data analysis and to apply natural language processing techniques to chart changes in the public's

awareness and understanding of family violence, in response to a high-profile public inquiry and related policy initiatives (McCosker et al. 2020). The work's impact centers on the collaboration processes that helped to re-shape the way the government evaluates policy outcomes, while also introducing new insights to associated agencies addressing the complex issue of family violence.

In Australia, as elsewhere in the world, there is evidence that family violence is endemic, and despite government and community health sector investment and effort, new responses are needed. Ongoing research into family violence continues to bring the issue into public view. This field of research has highlighted gender inequality and violence against women, revealing a host of related societal factors and issues including low education, low socioeconomic status, and the experience of childhood abuse, among other factors (MacGregor et al. 2019; Phillips and Vandenbroek 2014). Growing evidence of this kind led to policy action in Australia, initiated by a government which has the responsibility to address family violence and social services.

In 2014, the Victoria State Government launched a major Royal Commission investigation into family violence (2015–2016), followed up by the establishment of new family violence prevention policies. The term 'family violence' has been used in policy and research settings since at least the 1980s, but in public discourse the issue is often more commonly referred to as 'domestic violence'. Similarly, the notion of gendered violence has been used to refer to the underlying dynamics of family and domestic violence. The emphasis placed on family violence through the Victoria Government's terms of reference for the Royal Commission inquiry is fairly unique in policy contexts, offering the additional opportunity to chart, through the affordances of digital methods, the influence of this language and related policies over time.

Measurement, evaluation, and evidence-informed policy have been key drivers for governments seeking to partner with researchers and to expand the use of innovative data collection and analysis techniques. To this end, the Victoria Government's Department of Premier and Cabinet produced a framework of outcome indicators, presented alongside extensive recommendations for public- and community-sector reform (Department of Premier and Cabinet 2017). These indicators tended to reflect aspirations for change and thus have been considered difficult to measure, particularly those related to improved awareness, understanding and attitudes towards family violence in the community.

It is difficult and costly to regularly produce reliable evidence of changing community attitudes about significant social issues. The National

Community Attitudes to Violence against Women Survey (NCAS), administered by Australia's National Research Organization for Women's Safety (ANROWS), offers a long-running and detailed survey of one key factor underpinning family violence (Webster 2018). It has been used as a general reference point for understanding community attitudes to gender equality and gendered violence. The Department of Premier and Cabinet sought to supplement this and related research in its efforts to address family violence.

Social media data analysis, with an emphasis on capturing more direct responses to events, incidents, and issues, has been shown to be an effective complement to traditional data collection and the analysis of attitudes and public discourse (Dragiewicz and Burgess 2016). In addition, the growing field of natural language processing (NLP) research now offers new techniques and approaches to generating insights from large text datasets (Maier et al. 2018).

Drawing on digital methods and NLP, we aimed to identify and chart changes in the public conversation, as well as in the knowledge and awareness of and attitudes towards family violence and violence against women, by studying this 'public conversation' using public social media data.

How Did the Project and Partnerships Develop?

Collaboration with the Department of Premier and Cabinet (hereafter, the Department) and a group of stakeholder agencies was central to our approach. The yearlong project was funded directly by the Department, which oversees the implementation and review of policy and performance, coordinating activities that can help the government achieve its strategic objectives. Rather than via a traditional funding process, where a project is proposed and funded on the basis of merit or strategic need, our research institute was approached by the Department with the goal of exploring the potential to use innovative data and analysis to address an issue of importance for the government. The relationship that ensued was a more collaborative, interdisciplinary, and problem-led research process.

The project started in mid-2018 with discussions with a Department team. The team leaders were particularly inspired by the NESTA report "Governing with Collective Intelligence" (Saunders and Mulgan 2017). They sought to undertake a project exploring the use of repurposed or novel data to solve an information and policy challenge. From the start, it was clear that our research partners were looking for an issue for the government where the various stakeholder agencies and departments would be willing to come

together to address a data-driven challenge. Essentially, the partnership aimed to develop data analysis capability within government to improve the evaluation of policy outcomes (Farmer et al. 2023).

Because the focus and scope of the work had not been proposed and designed beforehand, our Department collaborators led a collaborative, co-design and scoping workshop, in which a range of participants from across government were invited. To formulate the research target, the Department stipulated that it should be in a non-controversial topic area (strategically important for the Government), there should be pre-existing good relationships between relevant agencies and departments, and the stakeholders involved should be interested in a novel data project. Follow-up planning workshops helped to narrow the work's focus to the exploration of new methods using novel community data sources and NLP analysis to address the complex and multi-stakeholder problem of family violence.

At that point, additional partners from within government and associated agencies dealing with family violence were brought together, forming an Advisory Group that provided important input and a source of 'ground truth' for the analysis and outputs. The group included representatives from:

- Women Victoria, a state government department promoting gender equality and women's leadership.
- Respect Victoria, an agency funded by, but independent of, state government, dedicated to the primary prevention of all forms of family violence and violence against women.
- Family Violence Branch, Department of Premier and Cabinet, Victorian Government.
- the Victorian Government agency leading the implementation of many of the government's family violence reforms.
- Behavioral Insights Unit, Department of Premier and Cabinet, Victorian Government.

In working with these stakeholders, certain issues and considerations had to be navigated. The Department had its own Behavioral Insights Unit that was analyzing data, so these staff members were involved from the beginning so as to avoid replicating the work they were already doing. Some of the stakeholders at Respect Victoria were involved as researchers with the independent organization ANROWS, working on more traditional studies of gender attitudes through the national survey (NCAS) and the analysis of media reporting of gendered violence (Sutherland et al. 2016). Exploring with the stakeholders what existing work had been done as well

as what gaps had remained helped to refine and strengthen the project. For example, through workshops and discussions, we clarified how the Behavioral Insights group undertook social media data analysis, and in response we developed and offered them new and alternative techniques.

Research Approach and Rationale

We worked with government and community partners to target outcomes relating to changes in public attitudes and awareness. We assessed changes by analyzing: (a) the public submissions that informed the new policy (to establish a baseline *vis-à-vis* core family violence issues) collected in 2015; and (b) public discussion through social media data (specifically, Twitter [now X]) and news media reporting, to understand how the public conversation changed in response to public policy from 2014 to 2018. We used Twitter data for its consistency over time, its publicness, and its ease of access.

In collaboration with the Department and project Advisory Group, an iterative approach was taken to data collection, analysis, and interpretation. With the Advisory Group, we identified a set of data sources that could provide insights into public discussions about family violence over the five-year study period, allowing year-to-year comparisons.

1. *Collaborative analysis strategy*: The Advisory Group met six times during the project. Early workshops helped to establish questions to pursue in the data analysis and to examine the timeline of policy events from 2014. As data were analyzed – and explored through subsequent workshops – the Advisory Group guided the understanding of family violence discourses and also gave feedback on findings and input to aid analysis.

The data sources and methods of analysis were as follows:

2. *Submissions to the Royal Commission public inquiry (2015)*: Of the 838 public submissions published on the Royal Commission website, we extracted a stratified sample of 105 submissions, including those by service organizations, representative (peak) bodies, networks and research institutions, and local governments and individuals. The submissions represent a cross section of ‘informed’ and ‘experiential’ accounts of family violence, drawn from the organizations, local governments, researchers addressing family violence, and individuals who submitted accounts of their firsthand experiences to the Royal Commission.

The public submissions constituted a proxy for the attitudes and topics discussed by an ‘informed public’ – that is, the diverse individuals, encompassing the community sector and services as well as government and research voices, who have experiences of family violence or work with victim survivors or perpetrators. We analyzed the whole sample through word frequency and thematic clustering using Pearson Coefficient Correlation analysis (Pearson’s r) and, with the sample, used qualitative content analysis to establish a baseline for the key policy dimensions framing family violence.

3. *Twitter corpus* (Jan. 2014–Dec. 2018): A corpus of 99,840 Twitter posts from 2,819 geographically dispersed Australian social media users over the five-year study period, including a wide range of public voices, reactions, and responses. While not everyone uses Twitter, there were many on the platform who used it to engage publicly with issues such as family violence. This dataset was chosen because it contains a wide and often competing array of voices and perspectives, and because it established responses in real time to issues and events of the day.

To identify topics in the Twitter dataset over the five-year timeframe, a sampling strategy was used, generating a maximum of 500 tweets per week for keywords centered around ‘family violence’ and ‘domestic violence’. To inform the timeline analysis, this sample was supplemented by extracting the ‘Twitter counts endpoint’, which returns the total tweet count at each timepoint. This enables quantifying of tweets beyond the 500-per-week sample over the study period. Australian geo-located, English-language Tweets were collected on 1 Jan 2014 and 30 Dec 2018.

We applied the topic modelling method Latent Dirichlet Allocation (LDA) to the Twitter posts for each year. The topic modelling process established a range of topic options, which were then reviewed by the research team, with input from the Advisory Group, to identify the most coherent and distinct topics, with the number of topics varying each year.

4. *News Media corpus* (Jan. 2014–Dec. 2018): A corpus of 11,451 news articles from Australian national and regional news sources (including newspapers, radio, and TV) was extracted using Media Cloud (<http://mediacloud.org>). News articles contribute to the framing of political and social issues and discussion and offer a point of comparison to the Twitter and Royal Commission submissions datasets. LDA topic modelling was applied to the news media corpus, with qualitative content analysis assisting in the labeling of topic clusters.

Natural language processing and qualitative content analysis techniques were supplemented with a timeline analysis of the Twitter dataset to identify

peaks across the five-year timeframe and to match these with known policy or public events. Named entity recognition was also used to identify key individuals and organizations and their prominence at different moments along the five-year timeframe. Importantly, we presented a wide range of data visualizations at stakeholder workshops, inviting collective interpretation and an iterative analysis process.

Findings and Collective Interpretation

Overall, the approach to analysis was informed by established theory in policy analysis, frame analysis, and sociolinguistics. These approaches address the formation of public social issues and account for the role of language and communication in framing, shaping, and contesting the parameters of those issues (Bacchi 2009; Vliegthart and Van Zoonen 2011). We reported a range of findings that helped identify longer-term changes in the way family violence was discussed and assisted in estimating the main effects of the Royal Commission and subsequent policy initiatives.

As noted, after the Royal Commission inquiry the Department worked to establish a set of Family Violence Outcome Indicators (Department of Premier and Cabinet 2017). Some of these indicators were measurable through existing government data. Several were related to the measurement of community awareness, understanding, and attitudes. By co-designing methods that used historical public social media and news datasets covering a five-year period, the analysis moved beyond the short-term analytics that the government's Behavioral Insights team had relied on to gauge community responses to policy. And it provided a far more accessible and efficient approach to the National Survey (NCAS).

One noteworthy outcome of our collaborative approach to analysis was that it led to the focused application of our findings to the Department's own family violence outcomes indicators framework. So while the insights themselves were important for what they said about the public's response to government interventions and policy on family violence, the work exerted another form of impact as well: it addressed the Department's capacity to provide evidence of policy outcomes through alternative data analysis techniques.

Responding to the goal of achieving an increase in public awareness of what constitutes family violence, we were able to detail the differences and synergies in the framing and focus of the 'expert' voices provided through public submissions to the Royal Commission, and accounts through news

and social media. On social media, people discuss family violence using their own terminology (referring more often to ‘abuse’) and increasingly probe its causes, contexts, and linked issues like gun violence. In contrast, the news media remains tied to a set of standard tropes – tending to portray family violence as extreme, violent, and involving policing. We conclude that discussion of family violence increased, but there were different constructions of family violence according to the data source.

The government’s goal of increasing understanding of the forms and impact of family violence by perpetrators was an outcome indicator that our social media analysis was uniquely placed to address (Department of Premier and Cabinet 2017). Social media discussion regularly emphasizes diverse voices of those with firsthand experience or those involved in advocacy and service-provision roles, accounting for the specific actions of (mainly) male perpetrators and the gendered context of family violence and the impacts of violence. The focus on men’s actions and forms of abuse became particularly prominent in 2016, 2017, and 2018. While news media shifted to incorporate more lived experience accounts into their stories, they tended to perpetuate the stereotype that family violence is public, very physically violent, and sensational.

Similarly, in contrast to policy documents, Royal Commission submissions, and news articles, our social media analysis was able to show the public’s nuances in its recognition of the impact of family violence on victim survivors, another key outcome indicator (Department of premier and Cabinet 2017). We surfaced a wider range of terminology about types of abuse through social media discussion (verbal, physical, manipulation, financial, bullying). When given the chance to speak publicly, victims and those speaking on their behalf express experiences as varied forms of abuse.

These and other key findings that detailed changing attitudes, understanding, and awareness showed clear differences distinguishing social media discussions and discourse, accounts submitted to the Royal Commission, and news media discourse. The analysis brought forth a timeline of Twitter post-density around key policy and Royal Commission events over the five-year period of analysis. We also identified key actors through social media and news media analysis.

Identifying Research Impact, Challenges, and Opportunities

Through this project and partnership, we were able to map the characteristics of the public discussion surrounding family violence over a period of

significant public policy reform between 2014–18. Our analysis filled gaps in the public survey work that informed the government’s understanding of public attitudes to family and gendered violence, which it relied on to provide evidence of policy outcomes. During and after the project, the Department was able to draw on our evidence to reshape the core domains in their Family Violence Outcome Framework, matching more closely with our analysis, focusing on Victim Survivors, Perpetrators, Prevention and System.

The findings were shared through a series of workshops at different stages of the research, involving our Department partner investigators and representatives from a range of agencies that address family violence. The workshops were a two-way conversation, presenting work in progress and findings, but also helping to refine the scope, direction, and points of focus. We also engaged closely throughout with the government’s Behavioral Insights group, an analytics team that monitors public interactions with the government across a wide range of media including social media platforms.

The work led to ongoing relationships and engagement with the stakeholder organizations and agencies. Our team was invited to showcase at the state government’s key knowledge transfer event for the public sector, an occasion where methods, insights, and processes could be shared with the public sector workforce. The work also led to a new funded research project dealing with accessing, integrating, and analyzing the government’s longitudinal datasets on Family Violence.

Collectively, we recognized that one of the main benefits of the project was the sharing of research methods and techniques in data analysis, offering new approaches to government policy evaluation. This was a transdisciplinary approach in the sense that it crossed disciplines and incorporated public sector input to address a complex social and policy problem. We have since conceptualized this process as one of collaborative data capability, in recognition that the benefits of digital and data-driven research are realized more effectively through sharing of skills and knowledge, as a form of “collaborative data action” (Farmer et al. 2023; McCosker et al. 2022). That is, although governments have greater access to data than ever before, the development of the capacity to analyze and make use of that data is still in its infancy. This is a role that interdisciplinary research teams can play.

The major lessons that came out of this project for our team were about the benefits of collaborative research despite its time-intensive nature. Public sector research partnerships can be challenging, and in large bureaucratic organizations require determined champions to drive experimentation and change. While we were fortunate to work with senior advocates within government, the project was hampered by staff changes, affecting continuity,

support and understanding of the work. Having contingencies and clearly documenting plans and progress are essential to manage this scenario. One of the benefits of public sector collaboration is the chance it offers to make a difference in the use of digital data for social good outcomes while contributing to oversight of ethical data practices in government.

We began the project without a fixed method, research questions or outputs, and while this can easily derail the traditional research process, we learned the benefits of adapting and collaboratively negotiating research with our public sector partners and domain specialists. For research students and early career researchers this can be a daunting prospect. However, we recommend finding like-minded collaborators across disciplinary divides to push the boundaries in the benefits of digital and data-driven research. Researchers can learn as much from our industry or public sector partners as they can from our knowledge base and research findings.

Acknowledgements:

This case study chapter is part of the project Social Outcomes of Policy: Helpful Intelligence and Analytics (SOPHIA), funded and supported by the Victorian Government Department of Premier and Cabinet. We acknowledge work of project members Jonathan Kingsley, Zoe Teh, Jasmine Knox, and Amir Aryani.

References

- Bacchi, Carol. 2009. *Analysing Policy*. Frenches Forrest, NSW: Pearson Higher Education AU.
- Department of Premier and Cabinet. 2017. *Family Violence Outcomes Framework and Indicators*. Melbourne: Victorian Government. <https://www.vic.gov.au/family-violence-outcomes-framework>.
- Dragiewicz, Molly and Jean Burgess. 2016. "Domestic Violence on #qanda: The 'Man' Question in Live Twitter Discussion on the Australian Broadcasting Corporation's Q&A." *Canadian Journal of Women and the Law*, 28 (1): 211–229. <https://doi.org/10.3138/cjwl.28.1.211>.
- Farmer, Jane, Anthony McCosker, Kath Albury, Amir Aryani. 2023. *Data for Social Good: Non-Profit Sector Data Projects*. London: Springer Nature.
- MacGregor, Jennifer CD., Casey L. Oliver, Barbara J. MacQuarrie and C. Nadine Wathen. 2019. "Intimate Partner Violence and Work: A Scoping Review of

- Published Research." *Trauma, Violence, & Abuse*, 22 (4): 717–727. <https://doi.org/10.1177/1524838019881746>.
- Maier, Daniel, et al. 2018. "Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology." *Communication Methods and Measures*, 12 (2–3): 93–118. <https://doi.org/10.1080/19312458.2018.1430754>.
- McCosker, Anthony, Jane Farmer and Arezou Soltani Panah. 2020. *Community Responses to Family Violence: Charting Policy Outcomes using Novel Data Sources, Text Mining and Topic Modelling*. Melbourne: Swinburne University of Technology. <https://apo.org.au/node/278041>.
- McCosker, Anthony, Xiaofang Yao, Kath Albury, Alexia Maddox, Jane Farmer and Julia Stoyanovich. 2022. "Developing Data Capability with Non-Profit Organisations Using Participatory Methods." *Big Data & Society* 9 (1). <https://doi.org/10.1177/20539517221099882>.
- Phillips, Jeanet and Penny Vandenbroek. 2014. *Domestic, Family and Sexual Violence in Australia: An Overview of the Issues*. Canberra: Parliamentary Library, Department of Parliamentary Services. http://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1415/ViolenceAust.
- Saunders, Tom and Geoff Mulgan. 2017. *Governing with Collective Intelligence*. London: Nesta. <https://www.nesta.org.uk/report/governing-with-collective-intelligence/>.
- Sutherland, Georgina et al. 2016. *Media Representations of Violence Against Women and Their Children: Key Findings and Future Directions*. Sydney: ANROWS. <https://www.anrows.org.au/publication/media-representations-of-violence-against-women-and-their-children-key-findings-and-future-directions/>.
- Vliegenthart, Rens and Liesbet van Zoonen. 2011. "Power to the Frame: Bringing Sociology Back to Frame Analysis." *European Journal of Communication*, 26 (2): 101–115. <https://doi.org/10.1177/0267323111404838>.
- Webster, Kim et al. 2018. *Australians' Attitudes to Violence Against Women and Gender Equality. Findings from the 2017 National Community Attitudes towards Violence against Women Survey (NCAS)*. Sydney: ANROWS. <https://ncas.anrows.org.au/wp-content/uploads/2018/11/NCAS-report-2018.pdf>.

About the Authors

Anthony McCosker is professor of digital media and communication at Swinburne University of Technology, Melbourne, and a Chief Investigator in the Australian Research Council Centre of Excellence for Automated Decision Making and Society. He researches the impact and uses of social media, data and new communication technologies.

Professor **Jane Farmer** is the Director of the Social Innovation Research Institute at Swinburne University of Technology, Melbourne. She has a distinguished track record of research in rural health services design, innovations in health workforce and technology, community engagement, codesign community initiatives and social enterprise.

Arezou Soltani Panah is Principal machine learning engineer at Atlassian and Adjunct Research Fellow with Swinburne University of Technology. She has a PhD in Computer Science from RMIT University and has been an active researcher in natural language processing, exploring new frontiers in translational research and policy informatics.

8. Data Against Femicide: The Process and Impact of Co-designing Digital Research Tools

Helena Suárez Val, Catherine D'Ignazio, and Silvana Fumega

Abstract

Data Against Femicide is an action-research collaboration that aims to foster an international community of practice around femicide data. In this chapter, we present two tools we have co-designed and piloted with activists to facilitate their work and alleviate the burdens of workflow. *Data Against Femicide Email Alerts* is an AI-based system that detects cases of femicide from media sources; *Data Against Femicide Highlighter* is a browser plug-in that helps activists scan media articles for specific data. By outlining the process and impacts of this work, we aim to motivate others to fully utilize feminist participatory approaches that place technology development and data science in the service of activists, movements, and changemakers who are making a difference in the world.

Keywords: Data activism; Latin America; Feminist HCI; Gender-related violence

Introduction

Data Against Femicide is an action-research collaboration aiming to make a difference in the work of practitioners who engage with femicide data. Femicide is a phenomenon defined as the killing or violent deaths of women that is directly related to their gender, occurring at the intersection of multiple forms of oppression and resistance (Lagarde y de los Ríos 2010; Segato 2010). Data Against Femicide was started in 2019 by Catherine

D'Ignazio, director of MIT's Data + Feminism Lab, Silvana Fumega, then head of research at the Latin American Open Data Initiative (ILDA), and Helena Suárez Val, founder of *Feminicidio Uruguay*.

Before we met, Catherine had already begun doing research about femicide data via Salguero's work (D'Ignazio and Klein 2020, 22–47); Silvana had led the development of a Latin American regional standard for femicide data collection (Fumega 2019b); and Helena, as part of her feminist activism, had been recording femicide data for Uruguay since 2015 and was researching these practices through a doctoral project. We connected through this shared endeavor and, on a call from our respective locations in Boston, Buenos Aires, and London, we decided to work together. In 2023, after working on the project as a collaborator, Isadora Cruxên, Lecturer in Business and Society at Queen Mary, University of London, joined the leadership team. Given our various personal and strategic interests and our knowledge about the community of practice around femicide data, we outlined three aims for our collaboration:

- To foster an international community of practice around femicide data.
- To develop tools to support the collection of femicide data from media sources.
- To support efforts to standardize the production of femicide data where appropriate.

Our project has since grown and developed around these three strands. To understand the work of activists producing femicide data, we have (so far) met and interviewed forty-one individual and group activists across the Americas. To help alleviate the emotional toll and intensive labor that activists must put into producing quality femicide data, we collaborated with seven monitoring projects across three countries to co-design and pilot two tools specifically tailored to activists sourcing data from mainstream and local news media. Finally, we created spaces for an emerging community of practice around femicide data to meet, learn and collaborate with one another. More than 1,500 people from all around the world have signed up for our public events and courses, which were offered from November 2020.

In this chapter, we share the process and impact of co-designing digital tools, focusing on the two tools we co-designed with activists: *Data Against Femicide Email Alerts*, a machine learning system that analyzes many thousands of media sources every day and sends an email listing probable cases of femicide; and the *Data Against Femicide Highlighter*, a browser

plug-in that helps activists scan media articles for specific data by automatically highlighting key terms. After elaborating on the importance of collecting data on femicide and the challenges associated with this work, we go into our design methodology and the tools that came out of this process. We conclude the chapter by reflecting on how feminist participatory approaches to technology development and data science can make a difference.

Femicide and Data

Since the 1990s, the term ‘femicide’ has been used by, has empowered, and has mobilized the feminist movement in Latin America (Bueno-Hansen 2010). The term recently regained visibility with hashtags such as #NiUnaMenos in Argentina in 2015 or #VivasNosQueremos in Mexico in 2019 (Chenou and Cepeda-Másmela 2019). In English-speaking countries, the term ‘femicide’ was first used in the 1970s but was not widely adopted (Radford and Russell 1992); it later inspired the Latin American translation of the term, which has recently become widespread in public discourse due to the work of feminist activists.

The production and circulation of data about cases of femicide has been and continues to be a feminist strategy to make this type of violence visible and to denounce it. Despite a call from the United Nations to establish a femicide watch in every country (UN OHCHR 2015), official femicide data are often inadequate or neglected, constituting what D’Ignazio and Klein (2020) conceptualize as “missing data”. Moreover, the collection of femicide data has not been standardized, making the phenomenon difficult to study across time and space (Fumega 2019a) and curtailing the ability to design evidence-based actions and public policy (Sivarajah et al. 2016; Buteau 2016). Activists have responded by learning to research, collect, visualize, and distribute femicide data. They record data about the events’ contexts, locations, forms of killing, relationship between perpetrators and victims, demographic details about them, and other variables; most of all, activists make femicide a “matter of care” (Puig de la Bellacasa 2017) by recording and remembering the names of women whose lives and dreams were violently ended by femicidal violence.

Some examples of this careful work include María Salguero’s sad and well-known map of femicide in Mexico; feminist observatories in Argentina by Mumalá, *Ahora Que Sí Nos Ven*, and *Mujeres de Negro*; the femicide accountability project *Women Count USA* by registered nurse Dawn Wilcox and *Black Femicide US* by Rosalind Page, also a nurse; *Transdados*,

data_labe's project aiming to combat the invisibility of trans and LGBTQ+ people in official data in Brazil; and the Sovereign Bodies Institute's database of Missing and Murdered Indigenous Women, Girls and Two Spirit people. Thus far, we have found more than 180 femicide data projects around the world (D'Ignazio 2023), driven by activists, journalists, or academics, working individually or as part of collectives or institutions, most doing the work for little or no remuneration.

Design Methodology

Drawing from the interdisciplinary experiences and the multiple situatedness of the lead collaborators – as activists, advocates, and researchers, and as members (to varying degrees) of the community of practice working with and researching femicide data – it made sense to adopt a feminist participatory action-research approach working across the North and South (Gatenby and Humphries 2000; see also Caretta and Riaño 2016) with a strong ethos of care and a focus on building relationships through research and practice. We wanted to learn and invite others to learn by working with others. Our ethos is also guided by the principles of data feminism, which call on us to examine and challenge power, elevate emotion and embodiment, rethink binaries and hierarchies, embrace pluralism, consider context, and make labor visible (D'Ignazio and Klein 2020). Our design methodology walks in the footsteps of feminist human computer interaction (HCI) work that aims to challenge cisheteropatriarchal paradigms in computing and contribute to liberatory design practices through pluralism, participation, advocacy, ecology, embodiment, and self-disclosure (Bardzell 2010).

Individuals and groups that monitor femicide do so in diverse contexts and languages; they use different definitions, categories, and taxonomies, and focus on regional or intersectional specificities. Nevertheless, their methods and experiences of producing femicide data share enough common ground to allow for the development of tools that can support their work across these differences. All the femicide activists we met have encountered major challenges in information retrieval and extraction, which involve a significant workload and considerable emotional labor, with daily encounters with accounts of violent deaths adding mental health burdens. Because femicide remains a contested term and not all its manifestations are addressed in existing legislation, femicide cases are often not explicitly reported as such by the media, which is the main source of information for most activists. Activists must read between the lines of accounts often

marked by ignorant, sexist, racist, and transphobic biases (to name a few) to ascertain whether the reported violent death of a woman was a feminicide. Furthermore, mainstream media do not cover all cases, so activists must seek out smaller, local news outlets, such as amateur blogs or social media accounts, which are less prominent or do not show up on popular search engine results. While some data collection and reporting projects are led by data journalists or by people with access to technologists by virtue of being based in an institution with such support (for example, a university), most activists making feminicide data are neither data scientists nor researchers. Only by teaching themselves have they acquired the collecting, recording, visualizing, and circulating data expertise that they now possess. This lack of access to formal training and technological support makes data collection and dissemination more laborious and time-consuming.

We saw an opportunity to create tools that could better support this form of labor that most often is unwaged or under-waged. The Data + Feminism Lab at MIT has a mission to apply data and computational methods to gender and racial equity work, making it well placed to take the lead in supporting feminicide data collection and dissemination practices. Also, because of our feminist methodological framework and commitment, it was critical that we collaboratively work with activists and value their contextual knowledge, labor, and expertise. We were clear at the outset that we wanted to avoid, as Haraway (1991) warned, creating a God-trick machine that would a) potentially miss and/or erase cases of feminicide due to lack of context, and therefore add to the multiplicity of “discordant data” (Suárez Val 2020) already in circulation in some countries; b) be rejected by activists as an imperialist or extractivist appropriation of their knowledges and techniques; and c) divert resources and attention from local and grounded actors towards a Northern ‘White savior’. The decision was therefore to collaborate with activists who might see the benefit in investing time to work with us to design tools that could make a difference in their day-to-day work.

We started by meeting individuals and groups to understand their practices and needs. We were able to count on the diversity of our team to conduct semi-structured interviews and qualitative analysis in their native languages, which facilitated our rapport and allowed people to speak without being hindered by translation. In parallel, we started working with two co-design partners: Helena, wearing two hats as the maker of *Feminicidio Uruguay* and one of the principal investigators in *Data Against Feminicide*, and Dawn Wilcox of *Women Count USA*. The co-design team was thus made up of these activists, the *Data Against Feminicide* leads, and members and students of the *Data + Feminism Lab*. Between June 2020



Figure 8.1. Ideas from co-design sessions for tools to support femicide data activism.

and February 2021, we ran six co-design sessions with Helena and Dawn, applying interactive prototyping and brainstorming tools and techniques to generate ideas for the creation of technologies that could support the work of monitoring femicide.

The sessions started out very open-ended, with the co-design team brainstorming tools that might be useful to femicide data activists. Using digital sticky notes, the participants came up with dozens of ideas which were grouped around different functionalities, for example: tools to detect and record femicides, tools to enable collaboration, tools to store and archive news articles and other primary source information, tools to visualize femicide data (figure 8.1). Out of the early sessions, two candidate tools emerged, and we prototyped and refined those in later sessions. Once the group was happy with the wireframes and prototypes, the Data + Feminism Lab developed and built a beta version of these tools between December 2020 and March 2021, and we then piloted both in English and Spanish. For this phase, we recruited groups from Argentina (Mumalá and Mujeres de Negro Rosario), Uruguay (Feminicidio Uruguay), and the United States (Sovereign Bodies Institute, Women Count USA, Black Femicide US, as well as a small non-profit organization that chooses to remain anonymous). Pilot participants filled out weekly surveys and joined two two-hour focus groups over a period of two months.

Since participatory design aims to work directly with end users as domain experts at every stage of the technology design and development process (Simonsen and Robertson 2013), we knew participation in the project would involve a significant investment of time for all involved, especially activists. We therefore chose to compensate participants financially to cover the time expended for interviews and for taking part in the co-design process. This remuneration, however, would never be sufficient, as the passion, knowledge, and care these activists have for femicide data, which they so generously shared with us, is priceless. We also extended our friendship and solidarity with the people and groups that collaborated with us and built ties that will continue beyond this project.

The Data Against Femicide Tools

The Data Against Femicide Highlighter

Most femicide data activists scan the news to ascertain whether the reported violent death of a woman matches the criteria for femicide, then

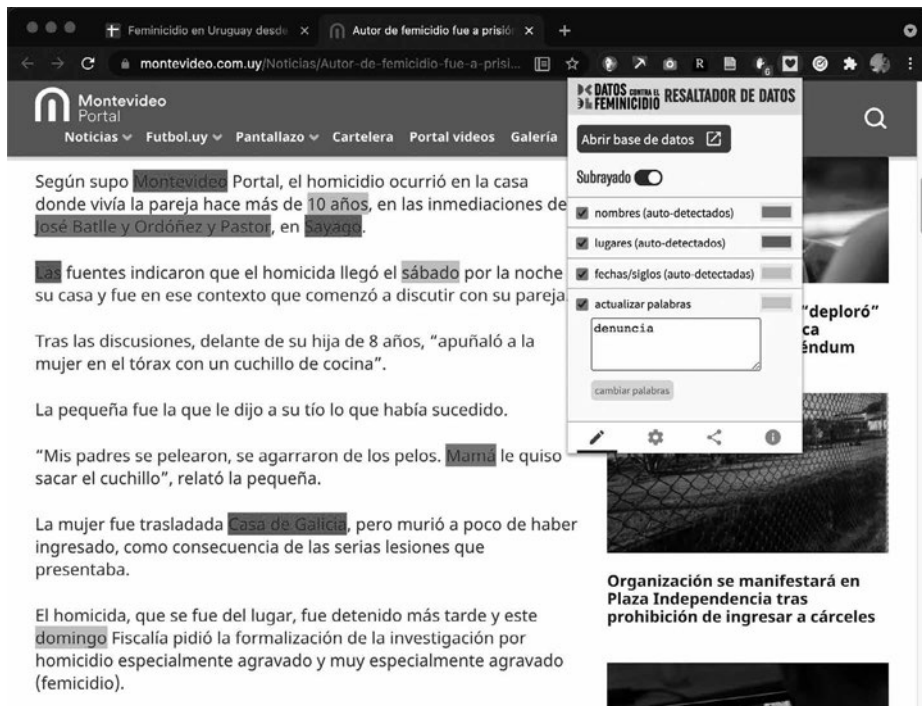


Figure 8.2. Data Against Femicide Highlighter.

manually extract data variables such as the names of the victims, location of the crime, and details about perpetrator from these news sources. To discover as much as they can about a particular case, activists will usually read between three to fifty articles about the case in order to triangulate the facts and collect as much information as possible. They often work with multiple open browser tabs or windows to copy/paste data, moving from the news articles to their database.

As seen in figure 8.2, the Data Against Femicide Highlighter is designed to facilitate this process. The tool is a browser extension for Chrome which auto-highlights on a webpage, with assorted colors, any words that represent proper names, places, dates, and numbers. Activists can also add custom words to search for and highlight. The Highlighter has a customizable 'Open Database' function that automatically arranges windows on their computer screen so that activists can see the news article and their own data-entry system on the side, top, or bottom of the screen. During the pilot phase, groups suggested enhancements, such as the ability to share an article with a colleague by email and a feature for highlighting custom words in assorted colors. These enhancements were implemented.

By visually highlighting key information in news articles and arranging the screen to suit the activists' preferences, making copy/paste faster, and by supporting in-team collaboration, the Highlighter reduced the overall time activists spent immersed in reading news articles about violence towards women.

The Data Against Femicide Email Alerts

To update the data at Femicidio Uruguay, Helena used to use Google Alerts to scan the web for new and old cases. She was often overwhelmed by the number of irrelevant and violent results she had to sift through. During our first meetings, we wondered whether we could develop a system to reduce the time spent searching for cases and the exposure to irrelevant news articles about violent events and their attendant emotional impacts. Our interviews with other activists confirmed that such a system could really make a difference, as their respective workflows also involved search engines and alerts; they reported similar feelings of dissatisfaction with the results for being too broad, outside their geographical focus, or matching the same case repeatedly, making it necessary to repetitively review the same articles or story.

Therefore, at the first co-design sessions with Dawn and Helena, we explored the idea of creating a 'better Google Alerts' system, where we imagined myriad ideas, from a fully automated concept where a tool would scan the news and enter results directly into a database, to a partially automated concept that would scan the news and alert activists the way Google does, but better. Because of the nuances of context and the need to read between the (biased) lines of media reporting, femicide detection in media is a complex intellectual and emotional task, and to avoid creating the God-trick machine, we leaned towards creating a partially automated system that would enhance the existing expert work of femicide data activists.

The Data Against Femicide Email Alerts tool uses a machine learning algorithm, a media database, and a mailing system to support femicide data activists in finding new cases and/or following the development of prior cases of femicide and gender-based violence. As in Google Alerts, activists can configure a search query, choose a geography of interest, and select how often they wish to receive emails. However, most importantly, the tool filters all results through a machine learning classifier specifically trained to predict the probability that an article is about a case of femicide. Moreover, the classifier groups multiple results from various sources *about the same case*, reducing the impact of receiving many unsorted articles and

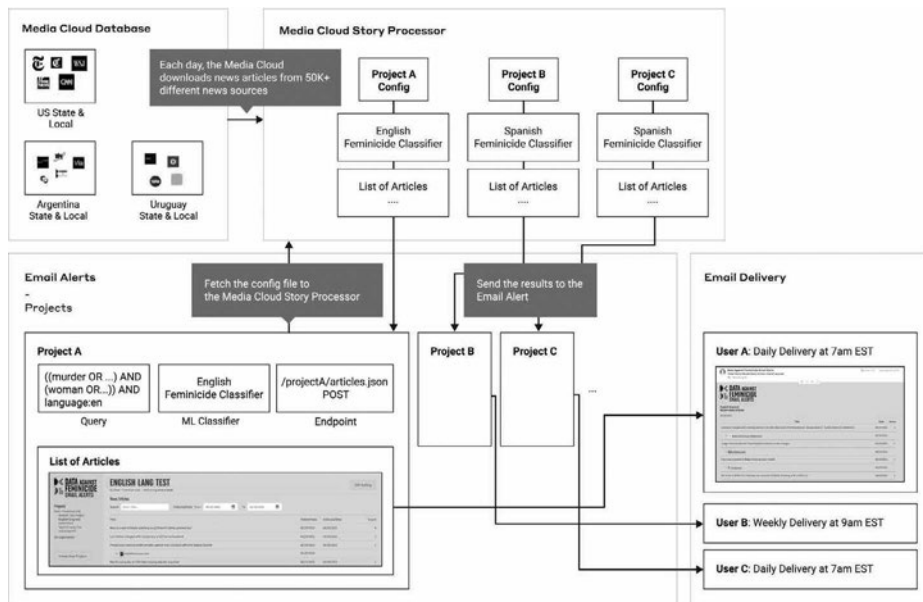


Figure 8.3. Data Against Feminicide Email Alerts system architecture. Graphic by Wonyoung So.

facilitating the activists' workflow. Finally, Data Against Feminicide Email Alerts, unlike Google, is not a black box. Because it relies on Media Cloud, an open-source platform for media analysis, it reveals its search sources, which include blogs as well as local and regional news outlets, and activists can suggest new sources. This transparency and flexibility are useful because activists often find local news and blogs more relevant than mainstream national media, and they constantly work to update their sources so as to overcome biases and the lack of reporting. At the same time, dependence on external infrastructure and systems can introduce vulnerabilities: Media Cloud's system outage in early 2022 temporarily prevented activists from using our tools.

While the Email Alerts worked well for groups who monitor all cases of femicide, the system proved less effective for groups that specifically monitor violence against Black women, Indigenous women, girls, and Two-Spirit people, and/or LGBTQ+ people. In many news articles, race, tribal status, gender identity, sexual orientation, and other key details are not mentioned at all or are wrongly ascribed. Moreover, these groups of people are often subject to underreporting or misreporting in the news media generally, especially when it comes to violence against women, where media reporting tends to focus on "worthy" victims (Stillman 2007). Groups and individuals who monitor cases of femicide for all groups, well aware of

these issues, try to overcome these distortions by acknowledging that their own works may reproduce the same biases (e.g., Helena calls Femicidio Uruguay “imperfect and incomplete”). But the biggest challenge, in terms of workload and emotional burden, is for those who belong to, and specifically monitor and care for cases that affect marginalized groups, who not only face the biggest gaps in terms of missing data but also endure violence disproportionately.

In line with our commitment to rethink binaries and hierarchies, and to avoid perpetrating erasure and missing data, the Data + Feminism Lab team continues to pilot this tool with two partners, experimenting together with tailored datasets and annotations. We hope to overcome these obstacles and improve the detection of cases of lethal gender-related violence involving people marginalized by sexism, racism, and colonialism.

Nonetheless, the Email Alerts system has proved to be an effective integration of a machine learning algorithm, a media database, and a mailing system into a tool that supports, rather than supplants, the expert labor, knowledge, and care of activists monitoring femicide. Pilot participants continue to use the tool and report that it has reduced their workload.

Conclusion

In this chapter, we briefly introduced what femicide data are and how this global phenomenon constitutes a case of missing data. To counter this missingness, activists have themselves produced femicide data, shouldering the attendant workload and emotional burden. We described how Data Against Femicide, an interdisciplinary collaboration among activists, academia, and civil society, supports and sustains the femicide data community of practice through various strategies, including using technology to develop tools that may facilitate activist efforts. We also presented the *Data Against Femicide Highlighter* and *Email Alerts* as concrete examples of how participatory design, guided by the principles of data feminism, can make a difference by not replicating technosolutionist and extractivist dynamics.

While we have not yet carried out a formal evaluation, we held an event in November 2021 where the pilot participants shared their experiences of working with these tools and other activists were invited to try them (ILDA 2021). At this event, activists told us the tools had helped with their work by reducing exposure to violent news and its resulting emotional toll; by facilitating case tracking; and by speeding up data entry. We also

imagined potential improvements for these tools, such as the addition of new sources, especially social media platforms, and the possible use of a common database format. Finally, activists raised the importance of the space the project created to reflect on their own processes and to meet others doing similar work. Rosalind Page, founder of Black Femicide US, said, “I do consider you sisters in the struggle”. For the research team, working on the different strands of this project has made us grow in our own feminist and data sociotechnical practices and research, and it has strengthened our resolve to continue pursuing this line of work.

Femicide is an urgent concern in a world where one in three women reports suffering some form of gender-related violence in her lifetime. More and better quality data are needed to understand and combat this “shadow pandemic” (UN Women 2020). While more tools or data on their own will never ‘solve’ problems of structural inequality, the Data Against Femicide tools suggest that small technological interventions, co-created in community, can indeed make a difference. How to secure adequate resources to keep the tools available for activists in the long term is an ongoing conversation for us. Significant student and staff labor to carry the work forward has been supported through funds made available by MIT or ILDA and its sponsors, but the project depends a great deal on the volunteer time and energy of its leadership team. Those who embark on technological developments in academia face the tricky issue of sustainability in a system that prioritizes and rewards innovation and invention over maintenance and code-fixing.

There are still many challenges to overcome in supporting the femicide data community of practice to continue researching, recording, and denouncing femicide, and in the face of these challenges we will continue to do so. We hope this chapter inspires other researchers and technologists to explore the full potential of feminist participatory approaches that place technology development and data science in the service of activists, movements, and changemakers who are already doing the work to make a difference in the world.

Acknowledgements:

We acknowledge all the individuals and groups who tenaciously take on the labor of producing and caring for femicide data and especially thank those who have generously shared their time and knowledge to contribute to Data Against Femicide.

Organizations that participated in the pilot

Black Femicide US, United States, <https://www.instagram.com/blackfemicideamerica/>

Feminicidio Uruguay, Uruguay, <https://femicididiouruguay.net>

Mumalá, Argentina, <https://www.mumala.ar/>

Mujeres de Negro Rosario, Argentina, <https://www.facebook.com/Mujeres-deNegroRosario/>

Women Count USA, United States, <https://womenscountusa.org>

Small nonprofit organization that focuses on Black women killed in police violence, United States

Sovereign Bodies Institute (SBI), United States, <https://www.sovereignbodies.org/mmiw-database>

Partner organizations

Datos Contra Feminicidio, <https://datoscontrafeminicidio.net>

Data + Feminism Lab (MIT), <https://dataplusfeminism.mit.edu>

Feminicidio Uruguay, <https://femicididiouruguay.net>

ILDA, <https://idatosabiertos.org>

Media Cloud, <https://mediacloud.org>

References

- Bardzell, Shaowen. 2010. "Feminist HCI: Taking Stock and Outlining an Agenda for Design." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1301–1310. CHI '10. New York, NY: ACM. <https://doi.org/10.1145/1753326.1753521>.
- Bernal Sarmiento, Camilo, Miguel Lorente Acosta, Françoise Roth, and Margarita Zambrano. 2014. *Modelo de protocolo latinoamericano de investigación de las muertes violentas de mujeres por razones de género (femicidio/feminicidio)*. ONU Mujeres; OACNUDH.
- Bueno-Hansen, Pascha. 2010. "Feminicidio. Making the Most of an 'Empowered Term'." In *Terrorizing Women: Femicide in the Americas*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, translated by Graciela Trevisan, 290–311. Durham: Duke University Press.
- Buteau, Sharon, Aurelie Larquemin and Jyoti Prasad Mukhopadhyay. 2016. "Open Data and Evidence-Based Socio-Economic Policy Research in India: An Overview." *The Journal of Community Informatics* 12 (2): 120–147.

- Canney, Julia. 2021. "Establishing a National Femicide Watch in the United States." *The Hill* (blog). 8 May 2021. <https://thehill.com/opinion/civil-rights/566544-establishing-a-national-femicide-watch-in-the-united-states>.
- Caretta, Martina Angela, and Yvonne Riaño. 2016. "Feminist Participatory Methodologies in Geography: Creating Spaces of Inclusion." *Qualitative Research* 16 (3): 258–266. <https://doi.org/10.1177/1468794116629575>.
- CEPAL. 2018. *Mujeres afrodescendientes en América Latina y el Caribe. Deudas de igualdad*. Santiago: CEPAL.
- CFOJA. n.d. "Femicide and the Law | Femicide in Canada". Canadian Femicide Observatory for Justice and Accountability. Accessed November 3, 2021. <https://www.femicideincanada.ca/courts>.
- Chenou, Jean-Marie, and Carolina Cepeda-Másmela. 2019. "#NiUnaMenos: Data Activism From the Global South". *Television & New Media* 20 (4): 396–411. <https://doi.org/10.1177/1527476419828995>.
- Deus, Alicia, and Diana Gonzáles. 2018. "Analysis of Femicide/Feminicide Legislation in Latin America and The Caribbean and a Proposal for a Model LAW." UN Women – United Nations Entity for Gender Equality and the Empowerment of Women. <https://lac.unwomen.org/en/digiteca/publicaciones/2018/12/analisis-legislacion-feminicidio-femicidio-modelo-de-ley>.
- D'Ignazio, Catherine. 2023. "Human-Centered Computing and Feminicide Counterdata Science." In *The Routledge International Handbook of Femicide and Feminicide*, edited by Myrna Dawson and Saide Mobayed Vega, 529–541. New York, NY: Routledge.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- D'Ignazio, Catherine, Helena Suárez Val, Silvana Fumega, Harini Suresh, Isadora Araujo Cruxên, Wonyoung So, María de los Ángeles Martínez, and Mariel García-Montes. 2020. 'Feminicide & Machine Learning: Detecting Gender-Based Violence to Strengthen Civil Sector Activism'. In *Mechanism Design for Social Good (MD4SG '20)*. online. <https://youtu.be/1EIDZ443wR4>.
- DuMonthier, Asha, Chandra Childers, and Jessica Milli. 2017. "The Status of Black Women in the United States." *IWPR* (blog). 7 June 2017. <https://iwpr.org/iwpr-issues/race-ethnicity-gender-and-economy/the-status-of-black-women-in-the-united-states/>.
- Fregoso, Rosa-Linda, and Cynthia Bejarano, eds. 2010. *Terrorizing Women: Feminicide in the Americas*. Durham, NC: Duke University Press.
- Fumega, Silvana. 2019a. "Open Data, Gender and Violence in Latin America." *European Public Mosaic* 7 (February): 36–47.
- Fumega, Silvana. 2019b. "Standardisation of Femicide Data Requires a Complex, Participatory Process, but Important Lessons Are Already Emerging |

- LSE Latin America and Caribbean". *LSE Latin America and Caribbean Blog* (blog). 8 August 2019. <https://blogs.lse.ac.uk/latamcaribbean/2019/08/08/standardisation-of-femicide-data-requires-a-complex-participatory-process-but-important-lessons-are-already-emerging/>.
- Gatenby, Bev, and Maria Humphries. 2000. "Feminist Participatory Action Research: Methodological and Ethical Issues." *Women's Studies International Forum* 23 (1): 89–105. [https://doi.org/10.1016/S0277-5395\(99\)00095-3](https://doi.org/10.1016/S0277-5395(99)00095-3).
- Haraway, Donna. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York, NY: Routledge.
- ILDA. 2021. "Data Against Femicide: Technology in Action." YouTube, December 15, 2021, 2:51:50. <https://www.youtube.com/watch?v=BhXG7pnzNW4>.
- Lagarde y de los Ríos, Marcela. 2010. "Preface: Feminist Keys for Understanding Femicide: Theoretical, Political and Legal Construction." In *Terrorizing Women: Femicide in the Americas*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, xi–xxv. Durham, NC: Duke University Press.
- Luján Pinelo, Aleida. 2018. "A Theoretical Approach to the Concept of Femi(ni)cide." *Philosophical Journal of Conflict and Violence* 2 (1): 41–63. <https://doi.org/10.22618/TP.PJCV.20182.1.171003>.
- Madrigal, Sonia, Ivonne Ramírez Ramírez, María Salguero, and Helena Suárez Val. 2019. "Monitoring, Recording, and Mapping Femicide – Experiences from Mexico and Uruguay." In *Femicide Volume XII: Living Victims of Femicide*, edited by Helen Hemblade and Helena Gabriel, 67–73. Vienna: ACUNS. http://femicide-watch.org/sites/default/files/Femicide%20XII_o.pdf#page=73.
- Puig de la Bellacasa, María. 2017. *Matters of Care: Speculative Ethics in More than Human Worlds*. Minneapolis, MN: University of Minnesota Press. <https://www.jstor.org/stable/10.5749/j.ctt1mmfspt>.
- Segato, Rita Laura. 2010. "Territory, Sovereignty, and Crimes of the Second State: The Writing on the Body of Murdered Women." In *Terrorizing Women*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, translated by Sara Koopman, 70–92. Durham, NC: Duke University Press. <https://doi.org/10.1215/9780822392644-005>.
- Simonsen, Jesper, and Toni Robertson. 2013. *Routledge International Handbook of Participatory Design*. Routledge International Handbooks. New York, NY: Routledge.
- Sivarajah, Uthayasankar, Vishanth J. P. Weerakkody, P. Waller, Habin Lee, Zahir Irani, Y. Choi, R. Morgan, and Y. Glikman. 2016. "The Role of E-Participation and Open Data in Evidence-Based Policy Decision Making in Local Government." <https://bradscholars.brad.ac.uk/handle/10454/14107>.
- Stillman, Sarah. 2007. "'The Missing White Girl Syndrome': Disappeared Women and Media Activism." *Gender & Development* 15 (3): 491–502. <https://doi.org/10.1080/13552070701630665>.

Suárez Val, Helena. 2020. "Datos Discordantes. Información Pública Sobre Femicidio En Uruguay." *Mundos Plurales* 7 (1): 53–78.

Unicef, UN Women, UNFPA, ILO, OSRSG/VAG. 2013. "Breaking the Silence on Violence against Indigenous Girls, Adolescents and Young Women." <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2013/5/violence-against-indigenous-women-and-girls.pdf?la=en&vs=1457>.

UN OHCHR. 2015. "UN rights expert calls all States to establish a 'Femicide Watch'." Press release. Geneva: UN OHCHR. 24 November 2015. <https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=16796&LangID=E>.

UN Women. 2020. "Infographic: The Shadow Pandemic – Violence Against Women and Girls and COVID-19". UN Women. 6 April 2020. <https://www.unwomen.org/digital-library/multimedia/2020/4/infographic-covid19-violence-against-women-and-girls>.

About the Authors

Helena Suárez Val is an activist, researcher, and producer focused on digital communications, data, human rights, and feminism. She has been recording and caring for feminicide data for Uruguay since 2015, at femicidiouruguay.net, and co-leads the international participatory action-research project, Data Against Feminicide.

Catherine D'Ignazio is a scholar, artist/designer and hacker mama who focuses on feminist technology, data literacy and data activism. She is the author of *Data Feminism* (2020), with Lauren F. Klein, and *Counting Feminicide* (2024). She co-leads the international participatory action-research project, Data Against Feminicide. D'Ignazio is an Associate Professor of Urban Science and Planning in the Department of Urban Studies and Planning at MIT and Director of the Data + Feminism Lab.

Silvana Fumega is a specialist on the intersection between data and inclusion. She is co-founder of Data Against Feminicide, Research Affiliate at the Data + Feminism Lab/MIT and member of The Data Tank's Global Advisory Council. She has served as consultant for numerous international organizations, governments, and civil society groups.

9. The Fairwork Project: Promoting Good Labor Practices in the Digital Platform Economy through Action Research

Tatiana López, Funda Ustek-Spilda, Patrick Feuerstein, Fabian Ferrari, and Mark Graham

Abstract

Though the digital platform economy in some contexts has created new employment opportunities, work in the platform economy has been widely criticized for its often precarious and exploitative character. One central mechanism through which digital platforms construct and conceal exploitative employment relations are information asymmetries among platforms, workers, and consumers. These information asymmetries are created through, among other things, a lack of transparency on how platforms allocate work, calculate payments, and use customer reviews for incentive structures and other rating-based work outcomes. Against this backdrop, the Fairwork project conducts action research to tackle these information asymmetries. In this chapter we introduce the Fairwork project and provide critical insights into our engagement with platforms as an action research strategy.

Keywords: Gig economy; Exploitation; Social transformation; Independence

Introduction

The digital platform economy has experienced significant growth in the past decade. An increasing number of services such as transport, delivery, and care-work are now arranged via digital labor platforms connecting

consumers with individual service providers. Although various typologies of digital labor platforms exist, we consider two broad types in our work: geographically tethered and cloudwork platforms. Geographically tethered platforms link workers with clients around specific localized activities such as cleaning, delivery, or transport; cloudwork platforms link clients and workers on a global scale around spatially dis-embedded activities carried out over the internet, such as software programming, design, data entry, translating, transcribing, or microwork (Woodcock and Graham 2020). Crucially, regardless of their geographical embeddedness or dis-embeddedness, digital platform companies play a central role in mediating the relationship between clients and workers – not only by providing the technological infrastructure but also by actively governing interactions between consumers and workers, e.g., through the management of job allocation and payment processes, the oversight of customer and worker complaint resolution procedures, and the establishment of work conditions (De Stefano 2015).

By facilitating labor market access, particularly to those who have traditionally faced higher entry barriers – e.g., women, migrants, and workers lacking formal qualifications – digital labor platforms have raised hopes for boosting employment especially in the Global South (see e.g., UNDP and FICCI 2021). At the same time, however, gig work remains often precarious and exploitative, characterized by contingent and insecure short-term work (Graham and Anwar 2019; Anwar and Graham 2021; Wood et al. 2019a). In this vein, Bertolini et al. (forthcoming) have highlighted the central role of digital labor platforms in creating new labor geographies that are shaped by three asymmetries between platforms and workers and that foster precarious work relations: contractual, financial, and informational asymmetries. Contractual asymmetries are constructed by platforms through practices of (mis-)classifying workers as independent contractors, thereby dis-embedding labor relations from national employment and social regulatory frameworks. As a result, many platform workers are deprived of benefits and protections such as pension plans or maternity leave. Financial asymmetries result from a platform's disproportionate access to capital: whereas precarious wages hinder workers' ability to save, most platform companies have access to venture capital and investments, enabling them to withstand economic slumps or disruptions.

Here, we aim to focus on the third type of asymmetries pointed out by Bertolini et al.: informational asymmetries. We argue that not only do platforms construct precarious labor relations shaped by informational asymmetries between platforms and workers, but they also create them with

consumers to conceal exploitative work relations (Rosenblat and Stark 2016; Ferrari and Graham 2021): On the one hand, platforms fashion informational asymmetries through limiting the information provided to workers and consumers on the organization of the labor and consumption process. On the other hand, platform companies collect and use data generated through each worker-consumer interaction to continuously enhance the company's market power and value capture from the labor process. As a result, in many cases neither clients nor workers have detailed information about the criteria according to which they are matched or about the fees and commissions that are charged. Moreover, platform companies monopolize information on overall supply and demand by withholding information from workers and consumers on the total number of jobs and workers on the platform. Given this lack of information, the work process becomes atomized and individualized, which keeps workers in a constant state of insecurity and constrains their capacities for collective organization (Cant 2019).

When criticizing the informational asymmetries on digital labor platforms that are used to create unfair and exploitative conditions, most studies point to the critical role of algorithms, which take on central management processes on most digital platforms (Rani and Furrer 2020; Wood et al. 2019b). We argue here that algorithms do not come from nowhere: they need to be actively programmed and fed by platform companies, which in turn have the power to create better or worse conditions for workers on their platforms. Precarious work is hence the result of direct choices and practices undertaken by platform companies rather than an inevitable structural feature of the gig economy harnessed by AI systems. This is also exemplified by the variations in working conditions on different labor platforms: whereas many platforms base their business model on the asymmetries mentioned above, there are also platforms that seek to create decent conditions for workers through the provision of permanent employment and other benefits (see e.g., Heeks et al. 2021).

Against this background, we introduce the Fairwork project, an international action research project committed to increasing transparency in the platform economy by highlighting the best and worst labor practices of the digital labor platforms. By evaluating platforms against five principles of fair work, the project aims to support workers and consumers in making informed decisions. Moreover, the project uses platform ratings as a means to create public pressure and incentives for platforms to actively improve working conditions.

In the following, we first introduce the Fairwork project and describe its specific approach to action research. Thereafter, we focus on the project's

engagement with platform companies as a central pathway for change within Fairwork's action research strategy. We provide insights into the challenges linked to engaging with platforms from the researcher's position and discuss strategies for addressing these challenges. We conclude with more general reflections on the role of action research projects like the Fairwork project in making a difference in datafied work settings.

The Fairwork Project: An Action Research Approach for Promoting Good Labor Practices in the Digital Platform Economy

Fairwork was founded in 2017 by researchers from the UK and South Africa with the aim of promoting fair work in the global platform economy. Initially operating in only two countries, the project rapidly expanded with teams in twenty-six countries by 2021. Fairwork activities, centrally coordinated by teams at the Oxford Internet Institute and the Berlin Social Science Center, have been generously supported by various funders, including but not limited to the UK Economic Social Research Council and the European Research Council, as well as the German Federal Ministry for Economic Cooperation and Development and the Ford Foundation.

Fairwork's action research approach is inspired by critical or radical research approaches that link knowledge production to the active promotion of social transformation (see e.g., Fuller and Kitchin 2007). Along these lines, Fairwork analyzes working conditions on digital labor platforms and uses this knowledge to engage with various non-academic stakeholders; its aim is to improve working conditions in the platform economy. These stakeholders are 1) gig workers and their organizations; 2) consumers; 3) policymakers and governments; and 4) the platform companies themselves. League tables – a ranking list of all platforms rated by Fairwork in a specific country or sector – are the central instruments Fairwork uses to engage with these stakeholders. By making transparent which platform companies invest in good conditions for their workers and which do not, Fairwork's league tables mitigate information asymmetries among platforms, workers, and consumers. Fairwork actively encourages stakeholders to use its league tables as resources for promoting positive social change in the platform economy: workers may use league tables, for example, to underpin their demands for better working conditions *vis-à-vis* platforms, whereas consumers and policymakers may rely on tables to make socially responsible decisions when using digital labor platforms. Fairwork furthermore uses league tables to create incentives for poorly performing platforms to improve their

ranking positions via the implementation of concrete changes for improving working conditions.

League tables show digital labor platforms’ performance in relation to the five principles of fair work developed by the Fairwork project in cooperation with workers, trade unions, state actors, researchers, and platform representatives. These principles are: fair pay, fair conditions, fair contracts, fair management, and fair representation. For each principle, a first and a second point are defined to rate platforms’ labor practices. Table 9.1 gives a summary of these principles and how they are operationalized for geographically tethered digital labor platforms.

Table 9.1. Operationalization of Fairwork’s location-based platform work principles. Source: Fairwork 2024

Fairwork Principle	First Point	Second Point
Fair Pay	Platform ensures workers earn at least the local minimum wage after costs.	Platform ensures workers earn at least a local living wage after costs.
Fair Conditions	Platform mitigates task-specific risks.	Platform ensures safe working conditions and a safety net.
Fair Contracts	Platform provides clear and transparent terms and conditions.	Platform ensures that no unfair contract terms are imposed.
Fair Management	Platform provides due process for decisions affecting workers.	Platform provides equity in the management process.
Fair Representation	Platform assures freedom of association and the expression of worker voice.	Platform supports democratic governance.

To evaluate a platform’s labor practices, Fairwork uses three research methods: desk research, interviews with workers, and interviews with platforms. Desk research involves the analysis of publicly available company documents as well as media and news reports on the platforms that have been chosen to be rated. Interviews with platform managers involve going through the Fairwork principles and requesting to see any evidence related to these principles and their particular thresholds. Interviews with workers involve conducting in-depth qualitative interviews (for geographically tethered platforms) and quantitative surveys (for cloudwork platforms).

Whereas we (and others) have discussed the various challenges and ethical implications when conducting research with gig workers elsewhere

(see e.g., Ustek Spilda et al. 2022), in this contribution we focus on our engagement with platform companies. Fairwork employs three main strategies to influence platforms' practices. First, Fairwork conducts large-scale media campaigns when new league tables are released, to raise broad public awareness about the labor practices of different platforms and thereby to mitigate information asymmetries between platforms and consumers. Second, Fairwork has launched the 'Fairwork Pledge,' offering corporate customers of digital platforms as well as NGOs, universities, and other civil society organizations the opportunity to make a public commitment to supporting fairer platform work by using platforms with better labor practices. Third, Fairwork also engages directly with platforms and, when invited to do so, provides feedback about how to improve the working conditions on their platforms. Hence, interviews with platform managements are not only engagements for data-collection purposes but are also spaces for discussion and direct engagement with the platforms: when evaluating the scores for a specific platform, research teams also develop detailed action plans for platforms to implement the Fairwork principles; these action plans are then discussed with the respective platform management teams.

Through its action research strategy, the Fairwork project has already made an impact during its first four years of existence. Following their engagements with Fairwork, several platforms have introduced changes to their business practices which have led to significant improvements for workers across the world, helping to mitigate contractual, financial, and informational asymmetries between platforms and workers. For instance, in line with the Fair Pay principle, the South African labor platform NoSweat committed itself to ensuring that workers earn more than a living wage in jobs secured via their platforms, thereby mitigating precarious work relations. Moreover, in line with the Fair Contracts principle, several platforms, including the Indian labor platform Urban Company, provided terms and conditions for workers in local and accessible language. In line with the Fair Management principle, various platforms such as the cloudwork platforms TranscribeMe and Workana have introduced institutionalized appeal mechanisms to allow workers to challenge unjustified disciplinary measures, and have introduced anti-discrimination policies for clients. Moreover, various platforms have made public commitments to accept and engage with collective worker bodies on their platforms, among them the South African labor platform SweepSouth. Most recently, the Spanish delivery platform Glovo, operating in more than 20 countries, developed a set of policies and social benefits for couriers in consultation with Fairwork, policies and benefits that will be rolled out in countries where Glovo operates.

However, pushing for social change by engaging with platform companies carries with it, as an action research strategy, several challenges for ensuring the independence of research and for maintaining a critical position towards platforms. In the next section we discuss these challenges and how we address them.

Challenges and Strategies for Maintaining the Independence of the Research Process

A central challenge for conducting action research in the gig economy that directly engages with platform companies lies in ensuring the independence of the research process and its findings. The independence and objectivity of research may be compromised, for example, when platforms fund research activities or are directly involved in the selection of workers for interviews. Against this background, we at Fairwork seek to preserve the independence of the research process and its results through three strategies. First, Fairwork finances its activities exclusively through third-party funding such as traditional research grants and grants for social transformation projects; none of its research or social transformation activities are funded directly by the platforms. Even when platforms solicit voluntary ratings and advice from the Fairwork project – as in the case of the above-mentioned cooperation with the food delivery platform Glovo – Fairwork does not accept any transfers of funds from platforms. Although this policy enables Fairwork to maintain its independence, the need for all research and engagement with platforms to be financed through third-party funding poses limits to the project's scope of engagement. Due to the limitations of personnel and its financial resources Fairwork has at times had to turn down requests for engagement and consultancy.

Second, to ensure the independence and objectivity of its research and to mitigate the potential risks for workers resulting from their participation in our studies, Fairwork tries to recruit participants independently from the platforms as often as possible. Only in exceptional cases do we ask platforms to distribute a call for participation in interviews or surveys via their mailing lists or community forums devoted to worker recruitment. We prefer to continue our communication with workers away of platform interfaces and move it to third-party sites, so that the platforms cannot identify workers who respond to our call (see Ustek-Spilda et al. 2022). Interested workers can then directly contact Fairwork researchers without the platform being involved in these interactions. When sharing workers' feedback with platforms in the rating process and making recommendations

for improvements to them, we ensure that all feedback is anonymized, and that no information is shared that would allow platforms to identify and/or single out individual workers or groups of workers.

Lastly, Fairwork implements several measures to maintain an independent and critical position when establishing collaborative communication with platforms in the rating process. To ensure that Fairwork ratings always accurately reflect actual working conditions on a specific platform, we conduct platform ratings annually so as to be able to grasp the improvements to, as well as the deterioration of, the working conditions on platforms. Treating platform ratings as a continuous process rather than a one-time interaction is also important, because Fairwork seeks to continuously adapt and strengthen its principles of fair work based on feedback from workers and other stakeholders. Hence, to maintain continued high ratings, platforms also need to continuously improve their working conditions. To that end, Fairwork always reserves the right to publicly criticize platforms for not meeting Fairwork's standards of fair work in the platform economy – independent of whatever sort of relationship Fairwork may have with a platform.

Conclusion

In this chapter, we introduced the Fairwork project, an action research project dedicated to promoting fairer work standards in the gig economy. Via the rating of digital platforms according to the five principles of fair work, Fairwork provides consumers, workers and other stakeholders with much-needed information about the working conditions on different platforms. In so doing, Fairwork aims to mitigate information asymmetries among platforms, consumers, and workers. Moreover, Fairwork creates impact with its annual platform ratings, as the rating process involves consulting with platforms to improve working conditions. A central challenge for engaging with platforms within the framework of an action research project is to maintain the independence of the research process and findings. The ability to acquire third-party funding for producing ratings, and for Fairwork's outreach activities, has been crucial to guarantee the independence of Fairwork's research so far and to maximize the impact of the project. Against this background, one central pathway for supporting social change towards fairer work in the platform economy is the expansion of funding opportunities for impactful action research. Government institutions and public agencies can play a central role in providing such opportunities by setting up funding programs dedicated specifically for promoting action research projects.

We hold that action research is a particularly apt strategy for promoting social change for two reasons. First, we need continued independent research to keep unravelling platforms' complex management and labor practices and to shed light on how these practices affect working conditions on a planetary scale. Second, research findings need to be linked to applied processes of engaging with platforms and other stakeholders for generating social change. Expanding those collaborations among researchers, platforms, workers, unions, and state actors is hence crucial for the development of alternative visions and practices for a fairer platform economy. Researchers interested in building collaborative relationships especially with workers and unions should, however, be sensitive to potential conflicts of interest that may arise when collaborating with a diverse set of stakeholders – including, for example, platform management and workers – at the same time. Transparent and open communication about the research process, the independent nature of the analysis, and the aims and boundaries of the project are therefore key for a successful collaboration between researchers and non-academic actors in action research projects.

Acknowledgements:

The Fairwork Project has received funding from the European Research Council (ERC), the Economic and Social Research Council through the Global Challenges Research Fund (GCRF, grant number ES/S00081X/1), the Federal Ministry for Economic Cooperation and Development (BMZ), commissioned by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), and the Ford Foundation.

References

- Anwar, Mohammad Amir, and Mark Graham. 2021. "Between a Rock and a Hard Place: Freedom, Flexibility, Precarity and Vulnerability in the Gig Economy in Africa." *Competition & Change* 25 (2): 237–258. <https://doi.org/10.1177/1024529420914473>.
- Bertolini, Alessio, Adam Badger, Fabian Ferrari, Mark Graham and Funda Ustek-Spilda. Forthcoming. "Reversing the Race to the Bottom in the Platform Economy: The Fairwork Approach." In *Handbook of Labour Geography*, edited by Andrew Herod.
- Cant, Callum. 2019. *Riding for Deliveroo: Resistance in the New Economy*. Cambridge: Polity Press.

- de Stefano, Valerio. 2015. *The Rise of the 'Just-in-Time Workforce': On-Demand Work, Crowd Work and Labour Protection in the 'Gig-Economy'*. ILO Conditions of Work and Employment Series No.71. Geneva: ILO.
- Fairwork. 2024. Location-based Platform Work Principles. Accessed February 14, 2024. <https://fair.work/en/fw/principles/fairwork-principles-location-based-work/>.
- Ferrari, Fabian, and Mark Graham. 2021. "Fissures in Algorithmic Power: Platforms, Code, and Contestation." *Cultural Studies* 35 (4-5): 814–832. <https://doi.org/10.1080/09502386.2021.1895250>.
- Fuller, Duncan, and Rob Kitchin, eds. 2004. *Radical Theory/critical Praxis: Making a Difference Beyond the Academy?* Vernon and Victoria: Praxis (e)Press.
- Graham, Mark, and Mohammad Amir Anwar. 2019. "The Global Gig Economy: Towards a Planetary Labour Market?" *First Monday*. <https://doi.org/10.5210/fm.v24i4.9913>.
- Heeks, Richard, Mark Graham, Paul Mungai, Jean-Paul van Belle, and Jamie Woodcock. 2021. "Systematic Evaluation of Gig Work Against Decent Work Standards: The Development and Application of the Fairwork Framework." *The Information Society* 37 (5): 267–286. <https://doi.org/10.1080/01972243.2021.1942356>.
- Rani, Uma, and Marianne Furrer. 2020. "Algorithmic Management of Work and Workers." *Competition & Change* 25 (2): 212–236. <https://doi.org/10.1177/1024529420905187>.
- Rosenblat, Alex, and Luke Stark. 2015. "Uber's Drivers: Information Asymmetries and Control in Dynamic Work." *International Journal of Communication* 10 (27): 3758–3784. <https://doi.org/10.2139/ssrn.2686227>.
- UNDP, and FICCI. 2021. "Impact of COVID-19 and Industry 4.0 on Future of Work for Women." https://www.in.undp.org/content/india/en/home/library/womens_empowerment/Impact_of_COVID19_and_Industry_on_Future_of_Work_for_Women.html.
- Ustek Spilda, Funda, Kelle Howson, Hannah Johnston, Alessio Bertolini, Patrick Feuerstein, Louise Bezuidenhout, Oğuz Alyanak, and Mark Graham. 2022. "Is Anonymity Dead?: Doing Critical Research on Digital Labour Platforms through Platform Interfaces." *Work Organisation, Labour & Globalisation* 16 (1): 72–87. <https://doi.org/10.13169/workorgalaboglob.16.1.0072>
- Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019a. "Networked but Commodified: The (Dis)Embeddedness of Digital Labour in the Gig Economy." *Sociology* 53 (5): 931–950. <https://doi.org/10.1177/0038038519828906>.
- Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019b. "Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy." *Work, Employment and Society* 33 (1): 56–75. <https://doi.org/10.1177/0950017018785616>.
- Woodcock, Jamie, and Mark Graham. 2020. *The Gig Economy: A Critical Introduction*. Cambridge: Polity Press.

About the Authors

Tatiana López is a post-doctoral researcher at the Department of Economic Geography at University of Würzburg, Germany. Priorly, she worked as a researcher in for Fairwork at the Berlin Social Science Centre. Her research focuses on transformations of work in the context of globalization and digitalization.

Funda Ustek-Spilda is a senior researcher and project manager at Fairwork, based at the Oxford Internet Institute; University of Oxford. She holds a D.Phil in Sociology from the University of Oxford. She writes on data justice, fairness and ethics with respect to emerging technologies regarding labour, gender and migration.

Patrick Feuerstein is a post-doctoral researcher for Fairwork at the Berlin Social Science Center (WZB). Before joining the Fairwork project, Patrick was a lecturer, research associate and acting professor of sociology with a focus on “Digitalization in the World of Work” and “Work, Organization and Economy” at the Georg-August-University in Göttingen.

Fabian Ferrari is a post-doctoral researcher at Utrecht University. As a member of the focus area Governing the Digital Society, he examines the governance of AI-generated media. Priorly, he was a researcher in the Fairwork project.

Mark Graham is the Professor of Internet Geography at the Oxford Internet Institute, University of Oxford. He is also Director of the Fairwork project. A full list of his publications is available at www.markgraham.space.

10. Advancing Equity through Data Practices: A Transformative Model for Organizational Change

Muna Osman and Hindia Mohamoud

Abstract

The Ottawa Local Immigrant Partnership (OLIP) is a multi-sectoral partnership of seventy-nine local organizations and community representatives promoting the equitable integration of immigrants in Ottawa. To remove inequities and enhance equitable policies and practices, organizations need a space to better understand inequities, develop inclusive data practices and secure infrastructure, and strengthen their capacity to collect and use disaggregated sociodemographic data. To address this gap, OLIP developed the Equity Ottawa initiative as a platform for knowledge mobilization, peer support, and shared learnings, as well as to promote accountability. This chapter outlines why and how we did this work, what challenges we faced, and the strategies we used to build more equitable and inclusive organizations in the city.

Keywords: Organizational change; Racial injustice; Ottawa; Immigrants

Situating Ourselves: Racism, Data, and the Role of Institutions

The global pandemic and racial awakening of 2020 shone a spotlight on issues of social and racial justice in Canada and around the world. Like other countries, Canada has a legacy of racism and oppression. This legacy is embedded in laws, enforced by government systems, and institutionalized across sectors. In our contemporary context, overtly oppressive systems have been more recently replaced by indirect and subtle forms of social exclusion

and marginalization. Despite changes in how racism is expressed, these mechanisms are deeply rooted in our public institutions through structures of governance, decision-making, organizational planning, data systems, and service delivery. Over the last decade, organizational leaders and staff are becoming more aware of the individual, interpersonal, institutional, and systemic manifestations of racism and oppression. This growing awareness has led many organizations and governments to reevaluate their role in sustaining the exclusion and marginalization of immigrant and racialized communities.

In Ottawa, close to 30 percent of the population are immigrants and about 33 percent racialized people (Statistics Canada 2023). Immigrants and racialized communities continue to experience discrimination, exclusion, and racism (Etowa et al. 2021; OLIP 2016; Williams et al. 2022). Disparities faced by immigrants span across many domains of life, including housing, education, health, policing, and employment. Immigrants and racialized people are also underrepresented in the labor force and leadership positions and are excluded from equitable access to public services, opportunities, and decision-making structures (Banting and Thompson 2021; Lightman and Gingrich 2018; OLIP 2016). Although these inequities are complex and persistent, organizations are actively looking for the best ways to leverage the insight and perspectives of diverse people in an equitable and inclusive way and to respond to these gaps and the changing demographics of the city.

Conversations about equity, both within and between organizations, have occurred against the backdrop of society's growing dependence on data to guide policies, decision-making, and planning. We know that quality data and information are essential for policies and services that offer effective solutions to the complex social, economic, and health realities that are products of systemic racism. Unfortunately, however, the data on the life conditions of immigrants and racialized people are limited. Existing datasets fail to disaggregate across key individual characteristics, such as race, ethnicity, immigration status, and gender among others (Choi et al. 2021). This results in outdated and fragmented information that renders the experiences of these communities invisible to policymakers when planning, allocating resources, and designing policy (Choi et al. 2021; Etowa et al. 2021).

Over the last decade, the Ottawa Local Immigration Partnership (OLIP) has served as a citywide platform to bridge the existing equity knowledge and capacity gaps as well as deepen organizations' responsiveness to community needs. Through the Equity Ottawa initiative, OLIP brought together

diverse partners, including data stewards from across the city, community representatives, academic researchers, equity specialists, nonprofit partners, and analysts from municipal and federal governments. This initiative used knowledge sharing, research, and community engagement to advance partners' understanding of inequity. Data systems and practices were central to this collaborative initiative. There was a need for population data on immigrants' experiences and for systems data from institutions on the effectiveness of structures and services to meet immigrant needs. These complementary sets of data are essential to understanding the historical and contemporary nature of disparities, to documenting community needs and systems initiatives, and monitoring progress and equitable change over time.

Effectively addressing these issues requires commitment, relevant knowledge, dedicated resources, and collective action (Banting and Thompson 2021; Kania et al. 2022; OLIP 2016). Although many organizations are committed to advancing equity, leaders lack the necessary expertise, data, and resources to address inequities in a meaningful way. Collective action – a collaborative effort by a group to achieve a common goal – can be a valuable and efficient way to share resources and address inequity (Kania et al. 2022). Through collective action, leaders can draw on the insights of organizations which have successfully been more welcoming and inclusive.

This chapter describes the exemplary case study of the Equity Ottawa initiative as an innovative way to support the processes of organizational change at the intersection of equity and the datafication of public institutions. We review this initiative's guiding principles, its successes and challenges, and the next steps to be taken. In this chapter, we highlight the complexity of equity issues in our datafied society. To guide this work, we take the position that while systemic oppressions inflict devastating impacts on immigrant and racialized communities, diverse communities continue to be resilient and resourceful, and their members strive for better life outcomes despite persistent experiences of discrimination and exclusion. An additional principle is that advancing equity is a two-way process, an acknowledgement that recognizes the reciprocal and reinforcing relationships between institutions and individuals. Institutions are mandated and aim to serve the public in a fair and equitable way; yet, it is the experiences of individuals that determine whether the services they provide are, in fact, equitable. Institutions and individuals must work together to understand the nature of inequities, including what policies and practices sustain them, and the actions and strategies needed to ensure fair, timely, and high-quality access.

Ottawa Local Immigration Partnership: A Platform for Change

Canada is a global leader in immigrant and refugee resettlement (Triadafilopoulos 2021). The Canadian economy also relies heavily on international immigration to remedy increasing labor and skill shortages. Although the successful integration of immigrants is a national priority, all immigrants settle in and integrate themselves into cities, neighborhoods, and communities, and thus, the process of integration unfolds at a local level. As a result, the local characteristics distinctly drive the nature and dynamics of immigrant integration, including its local culture, history, and public-sector infrastructure. As such, government leaders and policymakers saw the need to enhance cities' local capacity to address growing issues of service gaps and disparities, as well as improve immigrant outcomes at a municipal level. To address these concerns, a collaboration in 2008 between the Canadian federal and provincial governments developed the Local Immigration Partnership (LIP) initiative to build local capacity to support and plan for immigrant communities (IRCC 2017). Close to 100 LIPs are currently operating in municipalities and regions across the country. These partnerships engage organizations across multiple sectors to promote local policy change and long-term strategic planning to support immigrant reception, settlement, and integration.

The LIPs, representing a unique model, have made significant progress in the development and implementation of innovative multi-sectoral practices over the last decade. Many LIPs conduct local research on community needs, build the capacity and cultural competence of mainstream services, collaborate with provincial and municipal governments, and enhance information sharing and strategic planning for shared priorities across partner organizations. These strategies and notable achievements have increased the understanding of and coordination on issues specific to immigrants and racialized communities in cities across Canada (CIC 2013).

The Ottawa Local Immigration Partnership (OLIP) was one of the first established from the 35 LIPs in Ontario. OLIP quickly became a partnership of 60 organizations working collaboratively to advance a shared vision to create a vibrant, prosperous, and inclusive city that is strengthened by diversity and the contributions of immigrants. This partnership engages settlement agencies, municipal bodies, health service providers, social service agencies, school boards, universities and colleges, employment service providers, employer associations, immigrant civic organizations and community stakeholders. Partner organizations meet regularly to facilitate action on sector priorities, respond to community needs, and track their

ongoing progress across five sector tables: education; economic integration; health and well-being; language; and socio-civic integration. Working across sectors, OLIP promotes strategic alignment, catalyzes community planning, mobilizes research and knowledge, and spearheads collaborative action to respond to immigrant needs and remove persistent barriers. Given its shared vision and governance structure, OLIP is well-positioned to support multi-sectoral partners in their efforts towards equitable organizational change, specifically through the use of disaggregated socio-demographic data.

Socio-Demographic Data for Advancing Equity

On a global and a local level, information and human experiences are becoming more datafied. As such, our everyday experiences are increasingly quantified and captured through digital information and processing (Kitchin 2022). We are ever more reliant on data to drive decision-making, planning, and resource allocation. This contemporary process of datafication is broadly consequential for our lives and contains the potential for immense benefits as well as the risk of persistent harms. These consequences are closely examined across diverse disciplines including political science, economics, critical data studies, and law, and more recently they have incorporated decolonial and anti-racism perspectives (Kitchin 2022). Together, these perspectives critically explore the intersections of power, oppression, and data.

Socio-demographic data are essential to understand and expose the complexity of inequity and the diversity of perspectives and experiences (Etowa et al. 2021). However, in the context of historical and contemporary inequities, data systems are informed and driven by deeply rooted biases (D'Ignazio and Klein 2023). Understanding the role of data in perpetuating inequity is a multifaceted process. It requires examining our approach to data collection, analysis, interpretation, dissemination, and use through a critical lens of inequity (Kitchin, Lauriault, and McArdle 2017). This lens highlights the pitfalls of the collection and use of data to understand the experiences of marginalized communities, but also the harms of data misuse. Regrettably, without this lens data risks harming communities by perpetuating stereotypes and simplistic narratives that distort complex issues, and by reinforcing biases (D'Ignazio and Klein 2023).

Advancing equity through data requires us to grasp several key elements. We need foundational knowledge of how data is used to justify and maintain

harmful policies and decision-making practices, historically and in our contemporary context. We must comprehend how processes and decisions driving the datafication of society can exclude diverse perspectives and experiences. We need to make visible how systems of power and privilege can limit who can access or benefit from this datafication process and who can be harmed by it. Finally, we need to include community voices to best understand the risks associated with personal data use, issues of privacy that affect marginalized people, and the way these practices compromise the rights of diverse individuals. We keep these issues in mind as we explore the building of local capacity for the collection and use of disaggregated socio-demographic data.

Need for Local Data Capacity

Evidence indicates that immigrant and racialized communities experience disparities across multiple life domains, including health, education, employment, and civic-political participation. These disparities are also influenced by age, gender, abilities, socioeconomic status and other variables. To understand the complexity and nature of these inequities, OLIP highlighted the need for timely, granular, relevant, and local data. Given the intersecting nature of life domains, effectively addressing disparities requires a multi-sectoral and collaborative approach, in addition to quality data. Yet, the data available to organizations was limited, fragmented, and exclusionary. This work effectively addresses the social bias and inequities, and faulty representations inherent in the situated and contingent; it reproduces the oppressive systems and power structures that conceive, produce, manage, and analyze the data (Kitchin et al. 2017; Loukissas 2019). By working directly with organizations, OLIP highlights the need to address data settings as well as the data practices, assumptions, and systems (Loukissas 2019).

Through partner mobilization and knowledge sharing, OLIP addresses the need to build organizations' capacity to collect and use disaggregated data to inform community planning, equitable service provision and immigrant integration. Yet many organizations face resource constraints as well as knowledge and capacity gaps; they lack diversity in their workforce and leadership and are siloed in their efforts. To build local data capacity, OLIP relies on several strategic activities, including partner mobilization, research sharing, and community engagement. Partner mobilization focuses on regularly meetings with researchers, data practitioners, and community members to build a shared understanding of data needs and challenges. Research sharing includes compiling and sharing research evidence, hosting

presentations by data experts, and sharing partner experiences. Community voices are incorporated through dialogues and exemplary narratives. These activities involve supporting direct-service providers and training them to collect, analyze, and use disaggregated socio-demographic data to understand service use and user experiences, and to evaluate effectiveness and identify areas of improvement. Organizations need to balance external funder expectations and data requirements to demonstrate their achievements and internal data needs. These challenges often lead to data duplication and inefficiencies. In the end, data is a key tool to understand whether marginalized communities have services available to them that meet their needs and are specific to their context in the best possible way. Data can be shared across organizations to build knowledge based on effective services and programs to reduce disparities across life domains. These strategic activities and dedicated resources help the social services sector, non-profit, and other public organizations to strengthen their capacity to collect and use data to promote accountability and sustainability in advancing equity and mitigating disparities.

Beyond Disaggregation

Although disaggregation is valuable, it is only one way of using data to advance equity. There are several additional considerations. Focusing on equity will require the incorporation of community perspectives on data collection and use, along with community involvement in governance issues, using a strength-based approach, and the examination of the systemic nature of inequity through multiple levels of data. It is crucial to have community perspectives on how data can address that community's needs, as well as on what data needs to be collected, how data are used, and how their unique experiences are represented. As partners appreciated the importance of data, they needed knowledge about data management and governance practices. Governance includes the data standards, practices, laws, and regulations governing how organizations can use data. However, governance also includes more subtle ongoing processes and decisions about who can access and use the data, under what conditions, and for what purpose.

One added consideration is the importance of using a strength-based approach to disaggregated data. In efforts to understand inequities, researchers might only consider data points about deficits, challenges, and negative outcomes. Furthermore, service providers might use a deficit-based approach as a strategy to secure funding, support, or resources to address community

needs. However, using a deficit-based approach reinforces stereotypes and simplifies the narratives about communities, narrowing our understanding of deeper root causes. As an alternative, we need to use data to speak to strengths and assets in communities as they navigate discrimination and exclusion. Data can also explore the ways that being under resourced and underserved can negatively impact communities and perpetuate disparities. This approach will promote a deeper understanding of inequity, expand the narrative to include causes, and counter and reduce stereotypes.

With increased disaggregated data, we risk individualizing our interpretation of inequities. By focusing on the individual, the burden of changing the particular inequity (or inequalities) is placed entirely on the individual: systems are not held accountable. Using data to advance equity requires we understand the systemic nature of oppression and how systems, policies, and practices sustain disparities. As the experiences of immigrants and racialized people become more visible in data, rather than place this responsibility solely on communities and their members, we need to hold governments and organizations accountable so that these outcomes can be improved.

Case Study: Equity Ottawa

The Ottawa Local Immigration Partnership (OLIP), with local funding from foundations, launched Equity Ottawa in 2012 to help committed organizations explore the capacities they need to become more equitable and inclusive of immigrants and racialized communities, in order to better leverage community strengths, and best respond to their unique needs in a culturally informed way. Since August 2012, this initiative has advanced three distinct phases. Phase 1 (2012–2013) and Phase 2 (2013–2016) successfully engaged different mainstream organizations, developed a shared vision, and promoted knowledge sharing through a community of practice. Phase 3 (2016–2019) developed and refined a Theory of Transformational Change and a corresponding Collective Action Plan of seven interconnected, interdependent strategies and actions that support organizations in their endeavors to be more equitable and inclusive. Through these three phases, it was clear that organizations were becoming more aware of Ottawa's changing demographics, and the corresponding change in community needs. Organizations, appreciating the complex, multi-dimensional, and rigorous nature of equity work, were seeking resources and tools as well as equity expertise, and they aimed to learn from the experiences of other

partners and sectors that were further along in their journey to be more diverse and inclusive.

Equity Ottawa supports organizations through three mechanisms. The initiative helps to develop an enabling environment to support organizations' efforts in the service of equity. Partner organizations attend a supportive community of practice as a safe peer-learning space to discuss challenges. The initiative compiles and shares knowledge resources with organizations through learning events, tools, and equity specialists.

Research and expertise on equitable organizational change confirms that this process is an ongoing journey that requires active engagement across organizational domains and processes to focus on changing systems and internal cultures. For many leaders, this process is fulfilling and challenging. The benefits of this process include renewed energy and a rejuvenated commitment to learn from the knowledge and insights of diverse communities. In contrast, the process of change can be uncertain and unsettling for individuals and organizations, as it counters firmly held beliefs and introduces novel ways of working. Therefore, exploring issues of inequity through collaborative efforts need to be cognizant of the individual processes and structural components of equitable change. The structural components of this process form the basis for a critical examination of the systems and practices within organizations that enable the implementation of their respective mandates. Systems originate from and are informed by underlying cultures and traditions that guide practices, assumptions, and everyday functioning of the system. Working with organizations on equitable change requires a systematic re-examination of internal systems and practices to evaluate how they impact and respond to the realities of immigrant and racialized communities.

To help guide this work, this initiative conceptualized a Theory of Transformative Change and co-designed a Collective Action Plan with partner organizations. The emerging Theory of Transformative Change recognized that progress on equity requires a multifaceted and systemic approach through the key levers of change. Each of these levers plays a distinct role in advancing equitable change. These levers include organizations, community, the public, and the broader enabling environment and system in which organizations are embedded. As the project works directly with organizations, they are an important lever of equitable change. Equity Ottawa informs the functioning of this lever by providing peer support and cultivating equitable leadership, developing strategies for change across internal domains, sharing knowledge to acknowledge and address institutional racism, bias, and power relations, and building the capacity

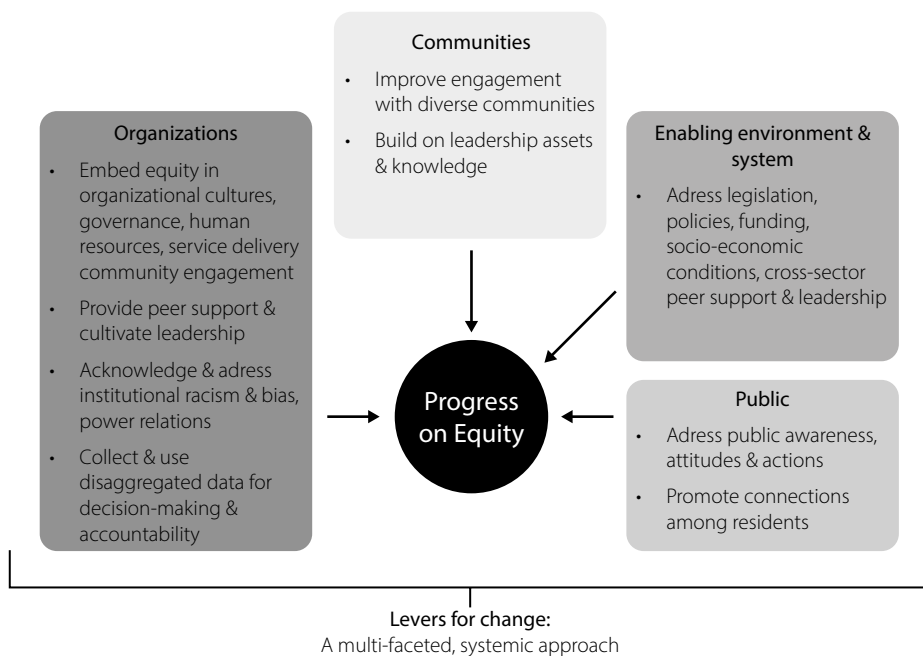


Figure 10.1. Equity Ottawa Theory of Transformative Change.

of organizations to collect and use disaggregated data for decision-making and accountability.

Communities are also a key lever of equitable change. Equity Ottawa aims to improve engagement with diverse communities and to build up the capacity of leadership and knowledge among community stakeholders. This project also engages the general public as a lever of change, through the building of public awareness, attitudes, and actions as well as the promotion of connections among residents. Lastly, organizations are embedded in a broader system of legislation, policies, and funding structures that can either hinder or enable their progress towards equity. As an enabling environment, this lever informs socioeconomic conditions and can enhance cross-sectoral peer support and leadership. Using this theory of organizational change, Equity Ottawa maps the levers and mechanisms of transformative change to eliminate institutional and systemic inequities.

The partnership co-designed a Collective Action Plan to outline the action and strategies needed to advance equity across domains. As a roadmap, this plan aims to support organizations in building a vibrant community that values diversity in identities and perspectives and supports the full inclusion of all residents. The actions outlined in the plan guide organizations in the strategic and meaningful efforts needed to navigate internal and external

dynamics in their individual journeys to address inequities and become more inclusive. In line with the theory of organizational change, this Collective Action Plan recognizes organizations are not isolated entities. Organizations are embedded in social, political, economic, and cultural circumstances that can either advance or hinder equity efforts. The plan also emphasizes the power and agency of organizations and their leadership to guide and inform legislation, policies, and social conditions. Additionally, organizations have the capacity to intentionally or unintentionally promote and perpetuate systemic disadvantage. Overall, this plan outlines the strategies and actions needed to support organizations to achieve their goals of equity and inclusion.

Understanding Equitable Organizational Change

Organizational change is a complex process. It involves identifying inequities and exclusionary practices across organizational domains, working with leadership to discuss, plan, and implement more equitable policies and practices, and monitor these changes over time to evaluate their impact on the organization and on community access and wellbeing. Building on the previous phases, the fourth phase (2020–2022) of Equity Ottawa worked with partners to determine how organizations can measure and report their progress on implementing anti-racism strategies and the actions outlined in the Equity Ottawa Collective Action Plan. This phase prioritized five organizational domains of change that often shape and reflect organizational values and systems of power. Given the embedded nature of inequity, organizations need to examine and re-examine internal policies and practices. These domains include human resources, governance and leadership, community engagement, service planning and delivery, and monitoring and accountability.

These domains are interconnected and interdependent. Efforts to change them can be captured as a continuum. This understanding of organizational change is based on evidence from psychological theories of behavioral change, business models of organizational change management, and the principles guiding Equity Ottawa. In addition to its five internal domains, Equity Ottawa also includes organizations building a foundation for equity in organizational culture and creating an enabling environment as two domains of change. Organizations can consider how they can build a foundation for equity through the sharing of knowledge resources across the organization, as well as through engagement with equity practitioners and via difficult conversations on how their practices can sustain inequity and

how they might improve these practices. Data is essential to the success of this change process. Partners are encouraged to collect and use disaggregated data to capture individual and systems level data. This would include data on multiple levels of their workforce, encompassing volunteers, staff, and senior leadership as well as data on service users. These data need to be both quantitative and qualitative to capture the people involved and the narratives about their experiences. Systems data can span practices in employment, governance, decision-making, service provision, and community engagement. Organizations were expected to collect these data internally and to do so regularly. Partners were also keen to document their progress on anti-racism strategies in their organizations.

Equity Ottawa also sought to collect data to monitor progress across the collective. Measuring progress required collecting data on diverse organizations within and across sectors. This was a difficult task given the different mandates and priorities of organizations, as well as the diversity of indicators and data systems used by organizations. A measurement continuum and standardized survey design were co-designed with partners to address this challenge. Through this collaboration a tool for collective monitoring was developed as a way for Equity Ottawa regularly to report on collective progress. The focus on monitoring progress recognizes the complexity of organizational change and works to measure indicators of progress throughout an organization. Monitoring the collective progress of partners requires an understanding of the mechanisms of organizational change and how equity was evolving or improving across partners and sectors. This approach is used to map a way forward for new and existing partners.

Equitable Change: Challenges and Successes

Over the years, Equity Ottawa has achieved many successes. This initiative has facilitated discussions and learning and knowledge events. It has sustained a regular community of practice for more than twenty-eight partner organizations and has included more than 100 stakeholders. Specifically, partner and other organization and community representatives engaged in webinars, conference presentations, and sector-wide meeting discussions. The Equity Ottawa team, with data science researchers, worked directly with individual partners to provide tailored feedback and training. Partners accessed information and resources from the OLIP website and received thematic meetings notes and resources through the Equity Ottawa email distribution list. Meetings provided a space and process for partners to explore concepts, articulate and analyze challenges, explore progress, and

exchange strategies through expert and partner presentations and facilitated discussions. Building awareness and knowledge was a key achievement of this initiative by engaging twelve external equity, anti-racism, data science, and organizational change experts, and connecting partners to local data experts to support data systems, collection, and use.

This project also came up against several challenges – among them, the diversity of organizations, the nature of the public sector, and changes among the partners – given the diversity of the partnerships and the sectors involved. Organizations varied in size, resources, structures, and mandate. Partners also have differing funding needs and challenges. This diversity can present challenges when it comes to developing resources and guidelines and supporting organizations to implement actions. Furthermore, the public sector faces challenges related to limited human, time, and financial resources, which impact the actions taken by the collective. The size of organizations can also give rise to specific difficulties. Smaller organizations may lack sufficient resources but are more nimble in implementing changes to their policies and practices. Larger organizations, though well resourced, can face barriers and resistance to influencing or changing policies and practices.

The nature of public institutions can also present challenges for this initiative. Public organizations can be willing and informed about what actions to take, but they may lack the resources required to implement and monitor these actions. Changes in resources and funding can compromise the scope and continuity of organizations' efforts. Additionally, public organizations can also experience turnover in leadership and staff. In addition, changes in the composition of the partner groups attending the community of practice also presented a challenge. Staff changes meant that new representatives of existing partner organizations were joining the sessions and new partners organizations were joining the Equity Ottawa initiative. As a result, the project team must support partners with varying levels of knowledge and experience. As new partners join, the group dynamics of the facilitated discussions, along with the concerns and needs of partners, can shift. As the partnership grows and experiences such shifts, the topics covered need to speak to the differing needs, interests, and concerns of partners organizations, which vary in size, mandate, nature, and structure.

The success and challenges of this initiative also could be intertwined. For example, partners reported having champions among senior leadership, staff, and/or community partners as key enabling elements of their work; however, limitations of human and financial resources, a key challenge, could diminish what they ideally would have dedicated to this work. Equity Ottawa provided the space to discuss these challenges, as partners made

reference to the valuable experience and insight they have accumulated over the years through their participation in this initiative. Overall, partners perceived the community of practice to be a valuable space, providing learning that they share with their individual organizations. In addition to the dialogue that they engaged in, partners appreciated having access to discussion papers, thematic summaries of dialogues, and resources on topics related to inequity, training, and assessment tools.

Next Steps

In addition to these main achievements, this initiative continues to grow and plans to expand on several fronts, including the number and diversity of its partners, the scope of the tools and resources it provides, and the specificity of the topics covered, which address new and emerging challenges. Progress is regularly monitored through an annual survey and interviews, and the assessment of the intermediate and long-term impact of actions taken by partner organizations on the lives of immigrants and racialized people.

We have seen partnerships grow exponentially over the years, both within partner organizations (by including more departments and representatives of a given organization) as well as by new organizations being welcomed to join the initiative. As organizations grow and continue to advance their actions and strategies, several challenges associated with data management and use remain to be addressed. In terms of data, the main next step for this initiative is to focus on socio-demographic data collection at the individual and system levels, to standardize data collection tools so as to promote comparability across partners, and to engage community members in issues of governance, data analysis, and interpretations, as well as data privacy and ownership.

At a population level, organizations require timely data to evaluate the levels of representation in leadership and their workforce, to better plan for community needs in service provision, and to ensure that disparities are reduced and eventually eliminated. However, many of the existing datasets at the national, provincial, municipal, and community levels are not disaggregated by individual characteristics such as age, gender, race, ethnicity, immigration status, or income, among others. Such gaps have serious consequences, impeding our efforts to measure, monitor, and evaluate equity. This lack of data can also have a cascading impact on planning, funding, and local disparities. Furthermore, organizations collect data in ways that reflect their mandates, resources, data capacity, and funding requirements, among other considerations. By supporting

partners in standardizing data collection tools, Equity Ottawa can report on partners' collective progress, map the changing demographics of specific institutions, and engage community members in guiding this process.

Additionally, due to systemic oppressions such as racism, many communities do not trust public institutions to collect personal data. Data can be perceived as bringing benefit and harm in contradictory ways. Being captured in data can make the experiences of inequity more visible, but it can also be threatening and expose diverse people to more discrimination and harm. To mitigate these risks, communities need to be reassured through increased transparency in the data process, as well as by the inclusion of their perspectives and by engagement with communities to incorporate their feedback in data use. Data are most critical for advancing equity when community perspectives are centered, and a strength-based approach is employed that highlights strengths and focuses on more than challenges and barriers. Using a strength-based and systemic lens on data concentrates on mechanisms promoting inequity in systems and practices rather than placing responsibility on the people targeted by oppression. Equity Ottawa provides a space for community members, equity experts, key stakeholders, and public institutions to discuss and reflect on these issues and to design a meaningful and collaborative way forward.

Conclusion

Data has great potential to enhance equity and social justice; however, data are not a panacea. This chapter outlines the case study of Equity Ottawa as an example of a multi-sectoral partnership bringing together academics, practitioners, and the social service sector to bridge the gap between research and practice in data science justice. This partnership was effective and mutually beneficial in strengthening local understanding of the situated and relational nature of data as well as in starting the journey to building capacity to collect and use disaggregated data. Working with local leaders, data stewards, equity specialists, and municipal-level data analysts, this initiative increased knowledge and awareness about inequities experienced by immigrant communities and the role of data in promoting equitable organizational change. Through collaborative knowledge sharing, data partnerships, and action planning, OLIP continues to work with a growing partnership to create a vibrant, prosperous, and inclusive city strengthened by diversity and contribution of immigrants through transparent and equitable data systems.

Acknowledgements:

The research for this chapter was supported by OLIP partners and generously funded by the Ottawa Community Foundation.

References

- Banting, Keith, and Debra Thompson. 2021. "The Puzzling Persistence of Racial Inequality in Canada." *Canadian Journal of Political Science/Revue canadienne de science politique* 54 (4): 870–891.
- Citizenship and Immigration Canada (CIC). 2013. *Local Immigration Partnership Handbook*. <http://p2pcanada.ca/wp-content/uploads/2013/02/CIC-Local-Immigration-Partnerships-Handbook.pdf>.
- D'Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Immigration, Refugees and Citizenship Canada (IRCC). 2017. *Evaluation of the Settlement Program*. Research and Evaluation Branch, IRCC. <https://www.canada.ca/content/dam/ircc/documents/pdf/english/evaluation/e2-2016-settlement-en.pdf>.
- Kania, John, Junious Williams, Paul Schmitz, Sheri Brady, Mark Kramer, and Jennifer Splansky Juster. 2022. "Centering Equity in Collective Impact." *Stanford Social Innovation Review* 20 (1): 38–45.
- Kitchin, Rob. 2022. *The Data Revolution: A Critical Approach to Big Data, Open Data, and Data Infrastructures*, 2nd edition. London: Sage.
- Kitchin, Rob, Tracey P. Lauriault, and Gavin McArdle. 2017. "Data and the City." In *Data and the City*, edited by Rob Kitchin, Tracey P. Lauriault, and Gavin McArdle, 1–13. Abingdon: Routledge.
- Lightman, Naomi, and Luann Good Gingrich. 2018. "Measuring Economic Exclusion for Racialized Minorities, Immigrants and Women in Canada: Results from 2000 and 2010." *Journal of Poverty* 22 (5):398–420. <https://doi.org/10.1080/10875549.2018.1460736>.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Mejias, Ulises A., and Nick Couldry. 2019. "Datafication." *Internet Policy Review* 8 (4). <https://doi.org/10.14763/2019.4.1428>.
- Mosley, Mark and Michael Brackett. 2010. *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK Guide)*, 1st edition. Bradley Beach, NJ: Technics Publications.
- Ottawa Local Immigration Partnership (OLIP). 2016. "Better Together: Cultivating Equitable and Inclusive Organizations in Ottawa: Summary Report on Phase

- I of Ottawa's Equity Project." <http://olip-plio.ca/wp-content/uploads/2013/07/OLIP-Equity-Report-EN-Low.pdf>.
- Statistics Canada. 2023. "Census Profile, 2021 Census of Population." Statistics Canada Catalogue no. 98-316-X2021001. Ottawa. Released November 15, 2023. <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E>.
- Triadafilopoulos, Triadafilos. 2021. "The Foundations, Limits, and Consequences of Immigration Exceptionalism in Canada." *American Review of Canadian Studies* 51 (1): 3–17.
- Walter, Maggie, Tahu Kukutai, Stephanie Russo Carroll, and Desi Rodriguez-Lonebear. 2020. *Indigenous Data Sovereignty and Policy*, 1st edition. Abingdon: Routledge. <https://doi.org/10.4324/9780429273957>.

About the Authors

Muna Osman is a research methodologist specializing in equity, inclusion, anti-racist approaches and broadly social and cross-cultural psychology. She has a PhD in Psychology from the University of Ottawa and has published in psychological journals on ethnocultural diversity, mental health, and youth civic engagement.

Hindia Mohamoud has more than twenty years of experience with building community solutions through research, partnership development, and program design. Until recently, she was the director of the Ottawa Local Immigration Partnership, where she is responsible for facilitating the implementation of community-wide vision and strategy for improving the settlement and integration of immigrants in Ottawa.

11. Advancing Critical Data Literacy through Justice-Focused Research: A Case Study of the Occupational Hazards of Mass Incarceration

Savannah Hunter, Lindsay Poirier, and Nicholas Shapiro

Abstract

This case study describes the model of multi-campus collaboration between Hack for California and Carceral Ecologies in advancing work on a project examining OSHA violations in US prisons, jails, and detention centers. The project – originally proposed by Carceral Ecologies – involved wrangling open government data that was notably fractured, inconsistent, and politically contentious. We conclude with a discussion of how facilitating multidisciplinary collaborations around social justice focused data science projects can advance critical data literacy education for students and produce more responsible and contextualized findings. After outlining the challenges involved in building and sustaining such collaborations, such as balancing educational opportunities and producing results that can lead to material outcomes, we chart some pathways forward.

Keywords: Critical data studies; Environmental justice; Prison-industrial complex; Pedagogy

Introduction

This case study outlines a collaborative effort to examine Occupational Safety and Health Administration (OSHA) violations in US prisons, jails, and detention centers through open government data. The project was initially conceived by Carceral Ecologies – a multidisciplinary environmental

research group at the University of California, Los Angeles (UCLA) – as one thread of their broader investigations into the environmental injustices of the prison-industrial complex. While other projects led by this group examined EPA violations in carceral facilities and the proximity of such facilities to toxic sites, this project centered OSHA data in recognition of the degree of scrutiny imposed upon indoor occupational facilities in monitoring environmental health risks as compared to similar settings for non-laboring people. In September 2019, Carceral Ecologies began collaborating with the University of California, Davis's (UC Davis) Hack for California – a group of data wranglers seeking to deepen understanding of social and environmental inequities through analysis of public data. In designing and advancing work on the project, collaborators came to recognize the project's reliance on open government data that were notably fractured, inconsistent, and politically contentious.

In this case study, we discuss the project's model of collaboration, along with how the project animated lessons from critical data and information studies (Dalton et al. 2016; Kitchin and Lauriault 2014; Iliadis and Russo 2016) for researchers engaged in the work. We discuss how collaborators learned to grapple with the tensions of presenting meaningful and actionable results from the data, while also acknowledging the context-dependence of claims made through data and the uncertainties woven through data analysis. As we reflect on the model of the multidisciplinary collaboration, we outline the challenges to building and sustaining such collaborations and chart some pathways forward.

Project Description: Background and Project Rationale

Hack for California was a research and learning cluster hosted in the UC Davis DataLab: Data Science and Informatics department from 2019 to 2021. The cluster convened multidisciplinary faculty, graduate and undergraduate students, and alumni to advance progress on data analysis projects seeking to deepen the understanding of social injustices facing Californians and beyond. Inspired by work in critical data studies (Dalton et al. 2016; Kitchin and Lauriault 2014; Iliadis and Russo 2016) Hack for California sought, as one of its underlying aims, to engage in data work both reflectively and reflexively. Participants were encouraged to critically examine the socio-political provenance of datasets engaged through the work and to assess the representational consequences of the judgement calls they made when cleaning and presenting data. The group met weekly to advance work on

a number of projects proposed by collaborating community organizations and research teams. One such team was Carceral Ecologies.

Carceral Ecologies is a multidisciplinary research group based in the Life Sciences division of UCLA, committed to investigating the environmental injustices of the prison-industrial complex. Due to the many failures of conditions-based prison reform at a local level, which have often enhanced the state's capacity to surveil, criminalize and incarcerate, the lab is attempting to approach the issue from a national perspective, querying the environmental health risks of the system as a whole. Drawing inspiration from Black (Davis 2003; Gilmore 2007; Kaba 2021), Latinx and Asian (Kim 2021), and Indigenous (CLEAR 2021) feminist practices, the lab seeks to query the institutional and structural violences of racialized mass incarceration and foster an inclusive research environment for those most impacted by the prison-industrial complex, who are often made to feel unwelcome in research. The lab is composed largely of undergraduate women of color, with many from the Underground Scholars Initiative (the UC formerly incarcerated and system-impacted student organization) and one currently incarcerated member who communicates via letter.

Collaboration goals centered around examining the environmental health risks of carceral facilities, drawing on government data. The collaborators gathered data from OSHA, a federal agency formed in 1970 under the OSH Act to set and enforce health and safety standards in private-sector workplaces. Key to the formulation of the project was the question of: who might emerge as unexpected allies in the quest for environmental justice for the millions of people incarcerated across the U.S? Could correctional officers or deputies also be endangered by the hazards of carceral environments and thus serve as at least fleeting *de facto* allies (Sobrinho 2021) in the quest for justice? As occupational indoor settings are much more tightly regulated than analogous environments for non-laboring bodies, we saw an opportunity to gain more insights about the dangers of the carceral built environment than we could glean via EPA data. We wondered whether, and if so, how the shared vulnerabilities of the human bodies that, unevenly, occupy toxic prisons might bring to bear the "queer productivity of toxicity" (Chen 2012, 211) in such a way that it might orient carceral laborers towards conversations on decarceration and searches for alternative rural economies. In line with these goals, the groups sought to: 1) obtain OSHA inspection and violations data for prisons, jails, and other carceral facilities; 2) explore the potential of OSHA data to lend insight into health hazards; and 3) examine carceral facilities' cumulative records.

The collaboration involved faculty and graduate and undergraduate students from Hack for California and the Carceral Ecologies Lab. The work

occurred over a period of two years, with team members shifting over time. Members worked at two-hour in-person working/coding meetings at UC Davis and Carceral Ecologies members called in via phone or video. During the COVID-19 pandemic all meetings moved to Zoom. Carceral Ecologies members played a critical role in orienting the project, interfacing with the US Department of Labor (DOL) at several stages for information and data access and conducting research on the data's variables and context. Hack for California members worked on project scoping, obtaining, wrangling, and analyzing data. Hack for California members were volunteers; however, the cluster received funding from the UC Davis DataLab to provide pizza at in-person coding meetings and to offer a small equipment stipend to a graduate student supporting the project. The Carceral Ecologies Lab members received either research apprenticeship credit or pay as part of this article's co-author Nicholas Shapiro's research start-up funds and a UCLA faculty senate grant.

Project Outcomes: Framing Critical Data Analysis Work

Engaging with OSHA data animated critical lessons from information and data studies for the group's members. Scholarship in these fields has critiqued metaphors that position data as a natural resource to be "mined" or "extracted" (Puschmann and Burgess 2014), calling attention to how such positioning obscures the human judgements and creativity invoked in the production, wrangling, and presentation of data. Contrary to dominant discourses that idealize neutrality in data work, ethnographic research into communities responsible for cleaning and analyzing data has shown that data analysts inevitably grapple with ethico-political concerns (Seaver 2021), even as various social forces and institutional structures render creative data labor invisible (Plantin 2018; Gray and Suri 2019; D'Ignazio and Klein 2020). Notably, this scholarship has shown that human discernment in data work is not limited to later analysis and visualization phases in a data science workflow; what is often referred to as 'raw' or 'pre-processed' data emerge as a result of judgments regarding 'what counts' (Martin and Lynch 2009), classificatory infrastructures that prioritize certain worldviews over others (Bowker and Star 1999), and messy social and environmental dynamics in data collection fields (Ribes and Jackson 2013; Biruk 2019). Thus, calls from these academic fields to engage data science work more critically and ethically have prompted data practitioners to investigate the cultural and political provenance of their data sources (Radín 2017; Loukissas 2019; Gebru

et al. 2020; Denton et al. 2020) and have suggested avenues through which such critical interventions can be woven into the data science curriculum (Dumit 2018; Poirier 2021).

Engaging Lessons of Critical Data and Information Studies

Throughout the collaboration, members engaged with lessons from critical data studies at every stage of the process, including obtaining, wrangling, and analyzing data. Grappling with these core concepts is critical for students who will go on to be data subjects, users, and analysts. We highlight a few examples.

Grappling with Data Semantics

The first step of the project involved obtaining Occupational Safety and Health Administration (OSHA) inspection and citation data for carceral facilities. Members decided to approach the objective broadly at first and obtain OSHA data for carceral facilities from all states since 1970. The Department of Labor (DOL) offers an API that allows users to obtain its data products; however, the API calls were not working due to a DOL technical error. The DOL also offers an interactive tool that enables users to obtain and subset enforcement data by clicking on different filters and downloading a dataset of the exact records desired. Members used the tool to filter the OSHA data to only include records with a Standard Industry Code (SIC) 9223, which is supposed to identify Correctional Institutions. While the tool quickly summarized the number of inspections, initial violations, employees exposed, penalties, and accidents, the group received an error message that they could not download the data because the file size was too large and were redirected to download data files that had not yet been filtered, aggregated, or analyzed from the DOL OSHA Enforcement Data Catalog. Through the manual effort to subset carceral facilities from this downloaded dataset, members realized that several carceral facilities in the dataset were missing SIC codes. Relying on SIC codes to identify carceral facilities in the data would delimit what would come to count (Martin and Lynch 2009) as a carceral facility in the data analysis. In response to this issue, members developed a dictionary of keywords often appearing in carceral facility names by which to additionally filter the data. While this identified more carceral facilities than the API or the interactive tool would have identified, it still posed limitations. The team acknowledged

that carceral facility names could be obscure (such as those appearing as acronyms) and undetectable via keywords alone. This practice prompted members to reflect on deeper questions: How do we define ‘what counts’ as a carceral facility? Why might there be inconsistencies between our own definitions and those of the agencies responsible for monitoring those facilities? Under what conditions are certain facilities unaccounted for?

Understanding the Social Dynamics Underlying Data Collection

After subsetting the data to carceral facilities, members grappled with how to focus the analysis. To scope the work, members filtered the data to investigations that were either closed or unresolved between 2010 and 2019. One co-author, having previous knowledge of OSHA jurisdictional issues, examined investigations by state and facility type (federal, state, and private). The OSH Act initially excluded state and federal employees from coverage. Almost half of states lack occupational safety and health coverage for federal or state employees, called “federal OSHA states” (OSHA, n.d.). In the data, we identified that seven federal OSHA states did not have any inspection records during the ten-year time period. An additional nine had no investigations for private or state-run facilities, although these states had investigated some federal facilities. Some states have OSHA-approved State Plans providing additional coverage for state/local government workers and/or stricter regulations for the private sector. Of those with state plans, there were thirteen with no investigations of private facilities during the time period. Members realized that while private-sector workforces were under the purview of federal OSHA, hardly any private carceral facilities appeared in the data. Additionally, given the lack of OSHA coverage for federal and state workers in many states (particularly the Midwest and South), it was hard to compare them to states with more comprehensive regulations covering private and public workers (like California).

Scoping the project involved learning the underlying social dynamics of data collection that influence what appeared in the records. Members realized that differences in OSHA jurisdiction meant that the scope of enforcement activities might vary, making potential comparisons between states challenging. Members learned that OSHA’s budgetary limitations meant that investigations were often motivated by complaints, leaving many potential violations uninvestigated (Weil and Pyles 2005). Additionally, prison workers are exempt from federal OSHA purview (Montoya-Barthelemy 2019). Finally, the Federal Bureau of Prisons operates a health and safety program for staff and prisoners, but the program’s data are not

publicly available. Considering these dynamics, the members focused the analysis on California's carceral facilities.

Politics of Interpretation and Presentation

During exploratory and descriptive data analysis, students began to understand the potential political implications of how they presented results. For example, cited facilities had records indicating an initial penalty followed by a current penalty that the facility actually paid. Some facilities received lower current penalties than had initially been assigned, and at least one facility had a citation marked as deleted. The students grappled with how to interpret changes in the penalty amount. Did a lowered penalty signal that the facility had successfully decreased their liability via OSHA's contestation process but the extent of the harm remained the same? The data provided no transparency regarding why a citation would result in a decreased penalty. The students learned that the data reflect a contested process in which facilities can challenge citations and penalty amounts. While data users can see the amount of monetary change, the reasons for the change remained opaque to them. Members grappled with how to think about and present these changes.

Similarly, members discussed how the interpretation of missing data could potentially reflect a particular politics. For example, only two of the 149 California carceral facilities investigated during the ten-year period listed any hazardous substances. Members struggled with how to interpret this finding. If no hazardous substances were listed, did this mean that the facility was free from hazard during the investigation or were hazards potentially unrecorded? The members learned that how they coded, interpreted, and discussed missingness could either present a picture of facilities performing well or could cast doubt on record-keeping.

Student Pedagogical Outcomes

Hack for California and Carceral Ecologies were explicitly focused on social justice projects that centered on marginalized communities. Engaging with lessons from critical data studies was vital to ensure that the projects remained true to those roots; for example, acknowledging that the data were partial and likely missed hazardous incidents that went unreported. Presenting missingness as ambiguity over the certainty that a prison was free from hazard was critical for students involved in the work as they reflected on the political implications of data analysis and presentation. In addition

to learning the skills of the research process (scoping a project, wrangling messy data, producing and presenting data analyses), the students learned, critically, the importance of examining data's context and of the careful and thoughtful interpretation of results. These myriad skills provided important professional development for the students involved, helping them to be not only expert data analyzers but also *critical* data analyzers – attuned to the sociopolitical contexts of data production and reflective of the ethico-political judgement calls made in their own data practice.

As a result of the project, members produced a Data Wiki documenting decision-making during data extraction and wrangling. The Wiki detailed different methods for obtaining OSHA data, described data quality issues encountered during cleaning steps such as standardizing facility names, documented how members decided to handle these issues, and considered how their re-coding might impact analysis. Additionally, it included a list of data definitions and resource links to help users gain a deeper understanding of how variables are defined and created by OSHA, as OSHA's Data Dictionary lacked deeper contextual information and even basic violation code identifiers that are not publicly available. The Wiki also identified potential variables for examining health hazards. The group produced reproducible code for extracting and cleaning OSHA data (Hunter et al. 2021) and a publicly available table demonstrating facilities' cumulative health and safety record for the variables identified. The table encouraged users to think about how they might conceptualize hazards by sorting and filtering the interactive columns. Finally, skills developed on the project helped one student to apply for and receive a two-year data science fellow position with a government agency.

Conclusion: Opportunities and Challenges

The interdisciplinary nature of the collaboration provided an important opportunity for students and the project as a whole. With the project convening diverse expertise in areas such as the carceral system, environmental health and justice, statistics, and computer science, skill sharing between members with different backgrounds often facilitated engagement with critical data studies topics, encouraging members to think beyond the confines of narrow disciplinary expertise.

Still, the collaboration faced a number of challenges. Hack for California and Carceral Ecologies operated according to different models. Carceral Ecologies was a research lab requiring students to work six hours per week in the lab. Hack for California followed a drop-in model, with volunteers

contributing to a project for two hours each week. Differing levels of commitment, turnover, and working time occasionally created barriers to project advancement. Future collaborations might intentionally engage in dialogue about how members will structure and pace their work together.

Many students joined both groups to learn research and data analysis skills. However, given the unstructured nature of the project, the messiness of the data, and the students' precursory experience with project design and implementation, members at times had difficulties scoping out milestones, organizing tasks, coding, and performing data analyses. While these struggles prompted important conversations and reflections on critical data studies topics, more extensive initial training could have supported more meaningful engagement. To this end, in its second year Hack for California, with contributions from Carceral Ecologies, ran a three-day workshop series on the basics of collaborating via GitHub, extracting open data via APIs, producing descriptive data visualizations, writing data documentation, and submitting public records requests. Requiring certain course prerequisites could have further supported the success of the collaboration and project.

Finally, the dual goals of advancing critical data literacy and producing results that examine and address carceral and environmental injustice at times became unbalanced, with pedagogical outcomes seemingly taking precedence over material outcomes. However, we came to realize that making a difference with data can take many forms. In this case, the most significant critical interventions involved documenting how social dynamics influence data collection, issues in data quality, and ambiguities in data missingness. Such documentation helps contextualize evidence and mobilize efforts to advocate for improved data infrastructure.

Acknowledgements

The research and collaboration for this project were supported by the University of California, Davis DataLab: Data Science and Informatics Department, and the University of California, Los Angeles Senate Faculty Grant. We would like to thank the Hack for California and Carceral Ecologies members, and specifically the following individuals for their contributions to the project: Selam Berekat, Konrad Franco, Jared Joseph, Amay Kharbanda, Ben Millam, Ambar Mishra, Ivy Molina, and Nathan Tran. We would also like to thank the following United States Department of Labor staff for their assistance accessing data: Giridhar Bandi, Kym Helmer, Patrick J. Kapust, and April Streets.

References

- Biruk, Cal. 2018. *Cooking Data: Culture and Politics in an African Research World*. Illustrated edition. Durham, NC: Duke University Press.
- Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Chen, Mel. 2012. *Animacies: Biopolitics, Racial Mattering, and Queer Affect*. Durham, NC: Duke University Press.
- CLEAR. 2021. "CLEAR Lab Book." CLEAR Lab Book (blog). 2021. <https://civiclaboratory.nl/clear-lab-book/>.
- Dalton, Craig M., Linnet Taylor, and Jim Thatcher. 2016. "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* 3 (1): 2053951716648346. <https://doi.org/10.1177/2053951716648346>.
- Davis, Angela Y. 2003. *Are Prisons Obsolete?* Open Media Book. New York, NY: Seven Stories.
- Denton, Emily, Alex Hanna, Razvan Amironesei, Andrew Smart, Hilary Nicole, and Morgan Klaus Scheuerman. 2020. "Bringing the People Back In: Contesting Benchmark Machine Learning Datasets." *Proceedings of ICML Workshop on Participatory Approaches to Machine Learning*. ArXiv:2007.07399 [Cs], July. <http://arxiv.org/abs/2007.07399>.
- Dumit, Joseph. 2018. "The Other Ninety per Cent: Thinking with Data Science, Creating Data Studies." In *Ethnography for a Data Saturated World*, edited by Hannah Knox and Dawn Nafus, 252–274. Manchester: Manchester University Press. <https://www.manchesterhive.com/view/9781526127600/9781526127600.00020.xml>.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2020. "Datasheets for Datasets." ArXiv:1803.09010 [Cs], March. <http://arxiv.org/abs/1803.09010>.
- Gilmore, Ruth Wilson. 2007. *Golden Gulag: Prisons, Surplus, Crisis, and Opposition in Globalizing California*. Berkeley, CA: University of California Press.
- Gray, Mary L., and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Illustrated edition. Boston: Houghton Mifflin Harcourt.
- Hunter, Savannah, Jared Joseph, Lindsay Poirier, and Nicholas Shapiro. 2021. "Carceral-Ecology OSHA Data Extraction and Cleaning Tool (version 0.0.2)." *Zenodo*. <https://doi.org/10.5281/zenodo.4825088>.
- Iliadis, Andrew, and Federica Russo. 2016. "Critical Data Studies: An Introduction." *Big Data & Society* 3 (2): 2053951716674238. <https://doi.org/10.1177/2053951716674238>.

- Kaba, Mariame. 2021. *We Do This 'til We Free Us: Abolitionist Organizing and Transforming Justice*. Abolitionist Papers Series. Chicago, IL: Haymarket Books.
- Kim, Nadia Y. 2021. *Refusing Death: Immigrant Women and the Fight for Environmental Justice in LA*. Stanford, CA: Stanford University Press.
- Kitchin, Rob, and Tracey P. Lauriault. 2014. "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work." SSRN Scholarly Paper ID 2474112. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2474112>.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Martin, Aryn, and Michael Lynch. 2009. "Counting Things and People: The Practices and Politics of Counting." *Social Problems* 56 (2): 243–266. <https://doi.org/10.1525/sp.2009.56.2.243>.
- Montoya-Barthelemy, Andre. 2019. "The Occupational Health of Prison Inmates: An Ignored Population and an Opportunity." *Journal of Occupational and Environmental Medicine* 61 (2): e74–76. <https://doi.org/10.1097/JOM.0000000000001514>.
- OSHA. n.d. "State Plans | Occupational Safety and Health Administration." Accessed November 5, 2021. <https://www.osha.gov/stateplans/>.
- Plantin, Jean-Christophe. 2019. "Data Cleaners for Pristine Datasets: Visibility and Invisibility of Data Processors in Social Science." *Science, Technology, & Human Values* 44 (1): 52–73. <https://doi.org/10.1177/0162243918781268>.
- Poirier, Lindsay. 2021. "Reading Datasets: Strategies for Interpreting the Politics of Data Signification." *Big Data & Society* 8 (2): 20539517211029320. <https://doi.org/10.1177/20539517211029322>.
- Puschmann, Cornelius, and Jean Burgess. 2014. "Metaphors of Big Data." *International Journal of Communication* 8 (0): 20.
- Radin, Joanna. 2017. "Digital Natives': How Medical and Indigenous Histories Matter for Big Data." *Osiris* 32 (1): 43–64. <https://doi.org/10.1086/693853>.
- Ribes, David, and Steven J Jackson. 2013. "Data Bite Man: The Work of Sustaining a Long-Term Study." In *"Raw Data" Is an Oxymoron*, edited by Lisa Gitelman, 147–66. Cambridge, MA: MIT Press.
- Seaver, Nick. 2021. "Care and Scale: Decorrelative Ethics in Algorithmic Recommendation." *Cultural Anthropology* 36 (3): 509–537. <https://doi.org/10.14506/ca36.3.11>.
- Sobrinho, Elena. "Environmental Justice and the Relational Politics of Abolition." Remote presentation, Annual Meeting of the Society for the Social Study of Science, Toronto, October 9, 2021.
- Weil, David, and Amanda Pyles. 2005. "Why Complain? Complaints, Compliance, and the Problem of Enforcement in the U.S. Workplace." *Comparative Labor Law & Policy Journal* 27 (1): 59–92.

About the Authors

Savannah Hunter is a sociologist of work and inequality. She studies how gaps in labor policy pose challenges for workers and their families. At the UC Berkeley Labor Center, Savannah contributes to research analyzing working conditions in low-wage industries and policies to raise labor standards.

Lindsay Poirier is a cultural anthropologist of data advocacy, governance, and infrastructure. Interlacing methods in cultural analysis and exploratory data analysis, Poirier's research critically examines social and infrastructural forces shaping how data becomes culturally meaningful. Poirier is the lead platform architect for the Platform for Experimental Collaborative Ethnography (PECE).

Nicholas Shapiro is a multidisciplinary environmental researcher and Director of Carceral Ecologies, a research group investigating the environmental injustices of the prison industrial complex. His current research focuses on the ways that science and technology have advanced or profited from mass incarceration in Los Angeles.

12. Empowering Citizenship Through Academic Practices: The Case Study of Amazonian Civic Media

Acilon H. Baptista Cavalcante and Ana Claudia Duarte Cardoso

Abstract

This chapter focuses on the potential contribution of academic practices to the training of young people for civic action in the periphery of an Amazonian metropolis. The text presents an account of the Data Firme project, carried out from 2018 to 2020 by a partnership of professors, technicians, and students from the Federal University of Pará. Data Firme used Design Thinking strategies to co-create media practices that disrupted the mainstream narrative, giving visibility to the area's cultural, social, and economic life and fostering civic movements with the support of digital media; such efforts form a central component of citizenship-building. This organized action resulted in social innovations being introduced in the relationship between community, media, and government.

Keywords: The Amazon; Civic media networks; Media practices; Informal settlements; Urban youth networks

Introduction

What difference can academia make in environments where there is social vulnerability? In addition to the obvious answers, based on scientific efforts aimed at solving problems – be they physical, social, or environmental in nature – this chapter presents a project that developed an unusual approach, strengthening a civic media network in areas with very little representation

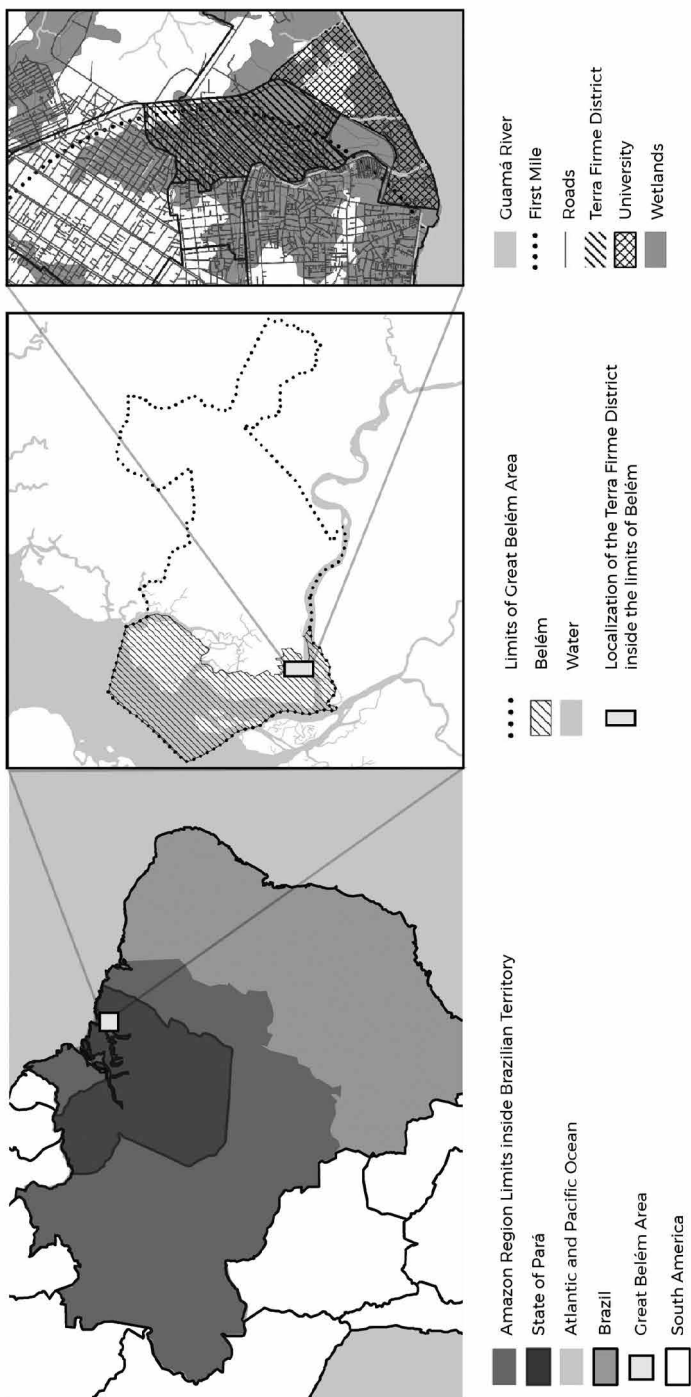


Figure 12.1. (1) Situation of the Great Belém Area inside Brazil. (2) Situation of the Terra Firme District inside the Great Belém Area. (3) The Terra Firme District near the University, highlighting the wetlands. Elaboration: Elis Brazil, source: www.ibge.gov.br.

in the city where they are located. This was a research and extension project developed at the Federal University of Pará, aimed at strengthening the population living on the periphery of Belém, the oldest and second largest city in the Brazilian Amazon (figure 12.1). The project, called Data Firme, worked with twelve young residents from Terra Firme, a district located in a flood-prone area that belonged to the Federal University of Pará; informally occupied, it was basically a landfill that had been built upon by the population. The project was led by young residents from Terra Firme who had entered university through the special access program created by the Quota Law (Federal Law 12711/2012), which enabled differentiated access for Black, Indigenous, and low-income populations, as well as people with physical disabilities, in order to increase the representation of these groups in public higher education.

The Data Firme project, which ran from 2018 to 2020, was financed by government funds and generated results that are just beginning to be felt by Terra Firme's population. These results will be analyzed through a consideration of transformations: those that took place in the relationship between the media and the community; in the government's relationship with the community; and in the social transformations triggered by training the young people who took part in the project.

The products delivered by the project were a web series, which registered the history and the living conditions and culture within the district, using testimonies from the residents themselves, and the creation of the Social Cartography of Terra Firme II, an action consisting of a series of activities that highlighted the daily problems experienced in the district due to a lack of sanitation, the precarious conditions of the streets, and the residents' lack of access to services.

Initially, this chapter adopts a socio-spatial-historical approach (Soja 2010) to characterize the Terra Firme district. It goes on to present the Civic Media Networks project, then ends with a reflection on the results and on how the academy can make a difference in contexts of social vulnerability.

Terra Firme: From Occupation to the Civic Media Networks

The occupation of the district occurred in a spontaneous and unplanned manner. In addition to there being neither organized road systems nor the provision of basic infrastructure, due to a land ownership impasse, a progressive densification took place. The occupation took over peoples' backyards and then spread to concrete slabs over the existing houses, which thus classifies

it as a precarious settlement. Towards the end of the 1960s, Terra Firme's population was 4,250 (Penteado 1967), but then, with all the transformations in the Amazon Region, the area would take in huge numbers of poor migrants, and by 1991 there were 59,231 residents in the district (Rodrigues 1996). In parallel, the city of Belém soared in population from 633,374 inhabitants in 1970 to 1,244,689 in 1991. Thus the city grew by 96 percent and the district experienced an impressive 1293 percent growth in its population.

Such intense urbanization is one of the main reasons given for the problems faced by residents today. The name Terra Firme has a certain irony since 83.75 percent of the district's 433 hectares are flood-prone areas (Pegado et al. 2014). During the 1980s, a period of major growth, the buildings and streets were largely made of wood, set on stilts and planks. There was no access to running water, nor was there any type of sewage system, and the streets were paved with solid waste (garbage and rubble) that came from the formal city (interview by Francisco Batista, 2019).

This type of occupation has occurred on the floodplains of the city, which are extremely flat areas intersected by small rivers, and have received large contingents of migrants from the interior, which was also undergoing an intensive restructuring process. These areas are made up of dense, informal neighborhoods that currently constitute the 'nearby' periphery (See figure 12.2) (Lima et. al. 2015, 161), given its proximity to formal and high-income districts within the current metropolitan center. There has been a trade-off between ease of access to public services and equipment for the migrants from that particular location and the inadequate conditions of the site for this type of occupation, which over three decades has resulted in the consolidation of this low-income district and in an improvement in the living conditions of the families that have settled in the area (Cardoso 2007, 83).

In addition to the physical precariousness and the land conflicts, one other factor has stigmatized the population, which is the labeling by public opinion of this area as an 'invasion'. This last aspect has been heavily played upon by the local media, which has portrayed this type of occupation from a perspective of marginality.

Canclini (2008) categorized such media practices as being part of a disinformation strategy, stating that "there are policies to distort and conceal information, such as the government and media strategies to concentrate and exclude large sectors of society, thus rendering them invisible" (Canclini 2008, 17). Indeed, this idea exposes the media's aim to create a homogenized idea of life in the city, thereby obscuring everything that deviates from ideal conceptions of a street, house, or district.

Over time, popular pressure has resulted in improvements in some of the social (education and healthcare services) and physical (water supply,

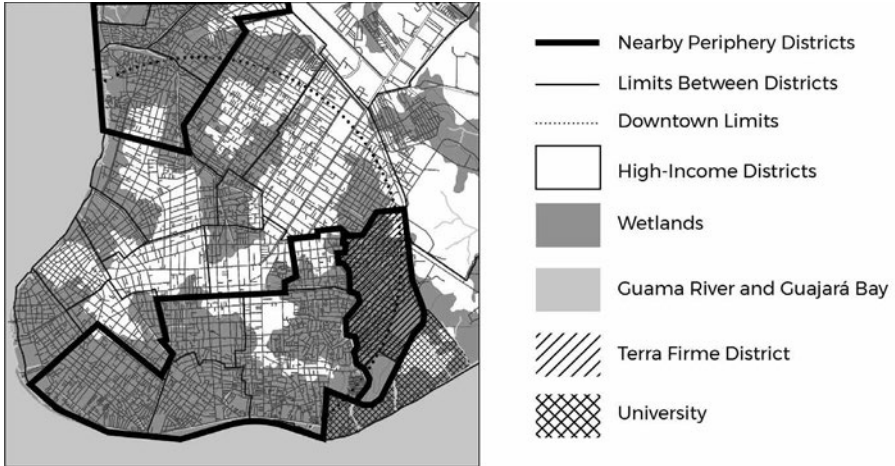


Figure 12.2. The nearby periphery districts inside the current metropolitan center. Elaboration: Elis Brazil, source: CTM/PMB, 2013.

sanitation, and drainage) infrastructures, and in the recognition of the area as a 'low-income district' (Cardoso 2007). However, this consolidation process has proved to be incomplete and the residents of the district currently live not only with the aforementioned problems but also with violence. During this decade, the low-income districts of the southern region of Belm have recorded the highest homicide rates and they suffer from the actions of militia groups and drug cartels.

During the 1990s, a type of media entertainment emerged in Brazil that explored urban violence on the periphery. It became known as community journalism but mainly worked with police news. Until the 2010s, Belm had five television programs, three radio programs, and two printed supplements dedicated to this type of content (Cavalcante 2020, 75).

A turning point came in 2014, after the so-called Belm Massacre. After the death of a militia police officer, one night of violence across nine peripheral districts saw eleven young people killed by a militia group as retribution. These sorts of 'responses' – actions to control the peripheral population through fear – have been taking place in Belm since 1996. With regard to public opinion, the media practices up to then have served to render this population invisible, as cited by Canclini (2008). From 2010 onwards, the stereotypical image of those living on the periphery as being marginal was so widespread that even these massacres' victims had resigned themselves to the violence they suffered. It was at this point that a change took place in media practices, which transformed the way the image of the periphery was portrayed.

Media practices are composed of meaning, competence, and materiality (Lunemborg and Raetzch 2018, 22), which are interconnected and function

through mutual reinforcement of the others. However, once the connections between the elements are broken, the practices attached to them also disappear, yielding new practices that re-emerge from what had existed. In the case in question, two changes contributed to the transformations. The first was the presence of a group of young people who, although they were born in the district, had access to a university education and so were much more aware of their role in society; they would reject the stereotypical 'marginalized' image that had been imposed onto them. The second change was to put an end to the monopoly on the narrative, which was achieved through the expansion of digital inclusion via access to smartphones and the internet. The expansion of telecommunication networks on the periphery occurred via the specific design of data packages intended for consumption by the low-income population. Among the metropolitan regions of Brazil, greater Belém is particularly exceptional for its 96.4 percent rate of connectivity, thanks mainly to the use of smartphones (Cavalcante 2020, 80).

Thus, fear gave way to indignation, which in turn led to action. Through digital media networks, the young people began to organize protests and to produce audio-visual content in order to raise public awareness. These actions made it possible to classify the groups that formed as civic media networks (Zuckerman 2014; Castells 2017); this emerging movement led to the proposal of the Data Firme project.

The Civic Media Networks in Terra Firme

Civic media networks, a global phenomenon, began around 2008 through movements such as Iceland's 'Kitchenware Revolution' and the Arab Spring (Castells 2017). They are characterized by a deep dissatisfaction among young people regarding the forms of participation offered to them in public debate; thus they become engaged in new forms of conducting politics (Zuckerman 2014; Castells 2017; Kahne et al. 2014).

After the Belém Massacre, young people used email groups to organize protests and debates, and through this mobilization, the leaders of the movement met and formed the first civic media group, Tela Firme (Interview with Ingrid Louzeiro, 2020). Tela Firme's performance led to a Parliamentary Commission of Inquiry investigation into the actions of the militia police, thereby bringing visibility to the problem in a way that reached academics and members of the arts community. However, public opinion and mainstream communication channels still needed to be reached, since media practices remained the same in the years that followed, even though

Tela Firme had kept producing content and sustained its activism through social media.

In 2017, talks were initiated to build a partnership between the university and the Tela Firme Group, propelled by the presence of some of the group in undergraduate courses, such as those that regarded arts, pedagogy, and multimedia projects. The main result of this partnership was to formalize the Data Firme extension project, presented below.

Data Firme

The project officially began in 2018 and was completed in May 2020. It emerged out of the Digital Language Incubator of the Faculty of Visual Arts, and a research project linked to the Postgraduate Program in Architecture and Urbanism (PPGAU), both from the same university.

The methodology adopted to build the ideas guiding the project was Design Thinking, which focuses on the construction of the process. In the words of Tim Brown, design thinking is “an exploratory process; done right, it will invariably make unexpected discoveries along the way” (Brown 2009, 15), which allows interaction with people in three stages: “an inspiration space, in which insights are gathered from every possible source; an ideation space, in which those insights are translated into ideas; and an implementation space, in which the best ideas are developed into a concrete, fully conceived plan of action” (Brown 2009, 63).

In this outline of action, the project defined the following axes: Promoting Citizenship, Increasing Income, and Constructing Social Capital. The table below presents the objectives generated within each of these three axes of action.

Table 12.1. The categories and impact desired by civic media groups when Data Firme began. These categories are grouped according to the Practical Guide to Generating Impacts – Artemisia – (2017)

Category	Objective
Promoting Citizenship	Increase the visibility of those in low-income and excluded groups. Physical access to the market for the sale of low-income production.
Increasing Income	Better market channels for small businesses.
Constructing Social Capital	Increase a sense of belonging. Increase in self-esteem related to presence in the neighborhood. Building a social network of trust, reciprocity, and cooperation with development.

After the objectives were agreed upon, the creation of the idea began: this was the stage where the products to be executed by the project were defined – namely, the seven episodes of the web series about Terra Firme and the Social Cartography of Terra Firme.

To set the action plan into motion, twelve undergraduate and postgraduate students were selected, all of whom lived in the district and were involved with civic media groups. In addition to the students, the project relied on five audio-visual professionals and two geographers, also residents of the district. Professors from the course on Technology in Multimedia Production also participated; supervision was provided by a PPGAU professor (who supervised the master's dissertation developed on the project); and another group from Terra Firme also took part: *Ame o Tucunduba* [Love the Tucunduba River], a collective of eight university students who promoted awareness-raising actions in the Tucunduba River basin.

The work plan activities were divided into four stages: planning, research, production, and post-production. Quantifiable goals were established for each stage, to be assessed at the end of the project. During the planning stage, the selected team was trained through workshops on scripting, planning, and Agile project management, and via meetings with university students. During the research phase, documentary research was conducted, interviews were filmed with residents from the first waves of occupation in the district (totaling more than 20 hours of recorded interviews), the Data Firme Hackathon was held by *Ame o Tucunduba* with the support of *Tela Firme*, and the II Social Cartography of Terra Firme, conducted by *Tela Firme* under the supervision of the geographer Francisco Batista, a leader from the district and the group.

The cartography fieldwork mobilized residents from the neighborhood, who collected information classified into six categories: Religion, Community and Cooperative Entities, Education, Commerce, Services, and Leisure, in addition to reviewing the databases referring to the road structure in the district (figure 12.3).

The cartography data were compiled into spreadsheets, which served as material for one of the videos in the web series, and were made available through digital networks, on flash drives, and in printed form. OpenStreetMap was used, which is a kind of Wikipedia of Maps (Meir 2015, 14) since there is an international community of contributors that feed it. OpenStreetMap was selected because it has advantages over other cartographic tools, such as ArcGIS, for this particular type of survey. It is online, it enables collaborative, shared contributions, it provides cloud storage, and it facilitates the development of applications for cell phones.

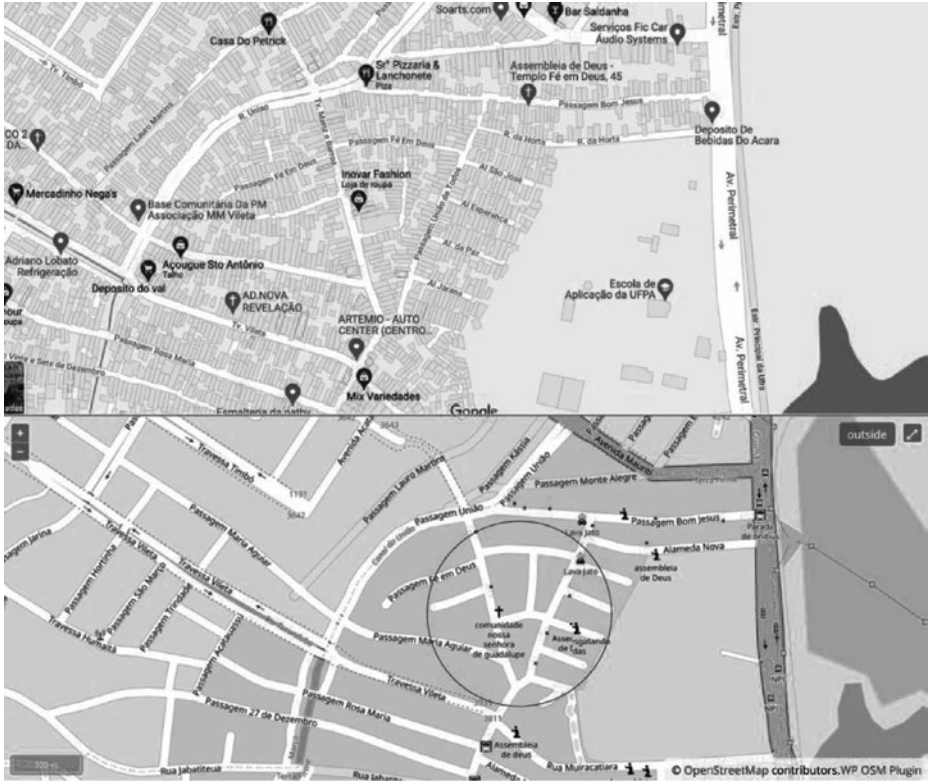


Figure 12.3. Comparison between Google Maps and Open Street Maps updated for the Social Cartography. The red circle shows missing streets in the Google platform. Using OSM icons, the project added churches, community centers, shops, and services to the digital platform. Author: Acilon Cavalcante, 2020. Source: <https://terrafirme.cartografiasocial.com.br/>.

At the end of the project, 300 boxes, with kits made up of printed maps and a flash drive with videos and a digital version of these maps, were shared with schools and with social actors to replicate the data across the community. The engagement achieved during the process was so intense that it attracted the attention of the city’s main communication vehicles and helped to bring about changes in media practices concerning the district.

What Difference Did Data Firme Make?

A project built in collaboration with the community runs up against many academic challenges. First, within the academic environment, there exists an outdated, mistaken, and elitist belief that positions the researcher on a higher plane than the public. It is not uncommon for teachers, however

well-intentioned, to believe they possess the answers to social problems. They frequently fail to appreciate the knowledge of residents concerning their own daily lives. On this particular point, one of the main merits of the project was that the civic media groups were the protagonists and the cooperation evolved according to their perceived needs.

Another challenge encountered was the issue of how to deal with technical problems. The work was conducted professionally, but it involved students who, for the most part, had never experienced working with multimedia projects in a job setting. Here, the professionals and teachers participating in the project had to take on a supervisory role, trusting that the results achieved represented the best possible outcome given the unavoidable limitations. Nevertheless, due to inexperience, the schedule was delayed, although the production quality was assured, thanks to the level of engagement of the students and of the Terra Firme community.

Data Firme contributed to the training of students, not only in the use of technological resources but also via their understanding of their role in the world. During the activities, other civic media groups emerged in the district that had relationships with the students working on the project (figure 12.4). Their activities increased their visibility, leading to changes in the relationship between the community and the media, and between these entities and the official discourse.

These changes were felt the moment the results were presented to the public. On October 16, 2020, the day the cartography products were launched at Terra Firme, there were radio interviews, television programs, and articles published in the city's print newspapers regarding the initiative. The visibility of the district increased in the mainstream media, a repercussion of the heightened awareness displayed by the youth groups on social media. In 2019, Tela Firme's fan page had 7,000 followers, and none of the posted content had reached more than 26,000 views. By May 2020, the page had reached 13,000 followers, and between January and May of that year, twelve posts attracted 17,000 to 600,000 views (figure 12.4). The post that achieved the greatest visibility was precisely the post that presented the party celebrating the entrance of young people from the neighborhood into the university.

The relationship between the public authorities and the district's population has also changed. As of 2019, Ter Paz, Territórios pela Paz [Be in Peace, Territories for Peace] became the state government's main public security 'pacification' program and a series of social programs were introduced around the peripheries. However, the program's organizers encountered great difficulties in engaging the population that was supposed to benefit

CIVIC MEDIA TIMELINE IN THE TERRA FIRME DISTRICT

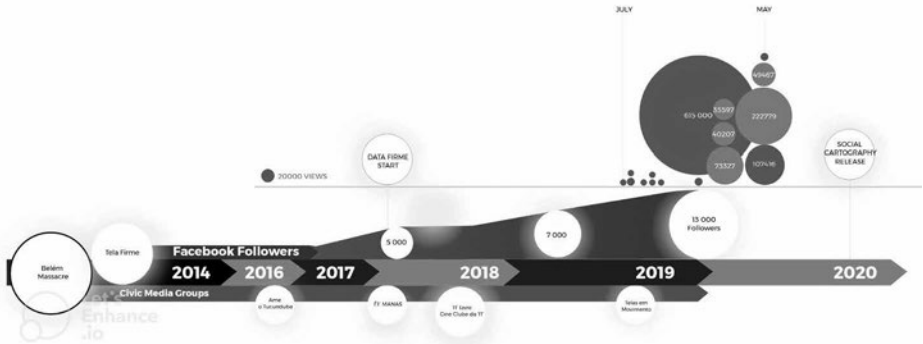


Figure 12.4. Timeline showing new civic media groups in the Terra Firme district since the Belém Massacre in 2014. It also shows the number of Facebook followers up to May 2020, and the visibility of Facebook posts reaching 20,000 views between July 2019 and May 2020. Before July 2019 there were no posts with 20,000 views. Author Acilon Cavalcante, 2020.

from it. They began with top-down strategies, selecting figures in the community who would replicate the proposals for the actions to be taken, thus persuading residents to participate in these social programs. In 2020, after the initiative had been running for two years, its results demonstrated that this strategy had not worked. It was then that the idea of a socially innovative intervention emerged at the Secretariat of Citizenship, and civic media groups such as Tela Firme started to be hired to promote Ter Paz within the communities.

In August 2021, Tela Firme was the first group to be hired and, using the collaborative networks it has established with other civic media groups, it has already operated in five districts, training 120 young people in actions related to citizenship.

Particularly outstanding among the results achieved by Data Firme is the training people have received based on this experience. Two students from the university joined, inspired by the project: Izabela Chaves in Cinema, and Walbster Martins in Multimedia Production. Both have pursued work in the area of human rights and have taken the lead, putting forth new visions of the world. Another student, Ingrid Louzeiro, was one of the founders of Tela Firme in 2014, and during the Data Firme period she began her master's degree. Not only was she one of the students most engaged with the community and the development of the work, but she also took her learning to City Hall, where she currently works developing an adult literacy program. This sample, of just three of the twelve students from the project, indicates the tendency of the reciprocal transformation of the young people and the

institution after the passage of the quota law, which is now at risk of being overturned by the current Brazilian government.

Conclusion

Educating young people is the main mission of the public university system in Brazil. It is essential, however, that higher education extends beyond merely producing professionals for the job market. Professionals should also be citizens who are aware of their role in the world. Access to the Federal University of Pará by the residents of the Terra Firme district, combined with the expansion of digital access, has begun a fast-track movement of social reengineering that has the potential to transform the lives of residents on the periphery. These transformations have re-presented these previously invisible individuals to the local society, with repercussions felt both in the media and in the official discourse. Nevertheless, there is still much to be done before the mission of promoting social justice and equal opportunities in the university will be achieved. Finding ways to confront media practices that exclude and enhance vulnerability was admittedly a small step. Still, it has been a significant step for the people living in districts like Terra Firme.

Acknowledgements:

This chapter is part of the “Characterization of the Civic Media Networks at the Terra Firme Neighborhood: Youth Engagement Seeking Citizenship through Digital Media” research project at PPGAU UFPA. The research was supported by a Nós na Tela Parliamentary Amendment and partnerships with Greenpeace, local entrepreneurs, and civic groups. We are also grateful to Francisco Batista, Ingrid Louzeiro, and all members of the Civic Media Movements from the Terra Firme district.

References

- Brown, Tim. 2009. *Change by Design*. New York, NY: Harper Business.
- Canclini, Nestor Garcia. 2008. “Imaginários Culturais da Cidade: Conhecimento/Espectáculo/Desconhecimento.” In *A Cultura pela Cidade*, organized by Teixeira Coelho, 15-32. São Paulo: Ed. Iluminuras: Itaú Cultural.

- Cardoso, Ana Cláudia Duarte. 2007. *O Espaço Alternativo: Vida e Forma Urbana nas Baixadas de Belém*. Belém: Editora Universitária da UFPA.
- Castells, Manuel. 2017. *Networks of Outrage and Hope: Social Movements in the Internet Age*. Cambridge: Polity Press.
- Kahne, Joseph, Ellen Middaugh and Danielle Allen. 2014. *Youth, New Media, and the Rise of Participatory Politics*. March 19, 2014. Oakland, CA: Youth and Participatory Politics Research Network. https://clalliance.org/wp-content/uploads/files/ypp_workinpapers_papero1_1.pdf.
- Lima, José Júlio Ferreira, Ricardo Bruno Santos, Lucinda Freitas de Assis Sena, Camila Leandra Araújo. 2015. "Estrutura Social e Organização Social da Região Metropolitana de Belém." In *Belém: Transformações na ordem Urbana*, edited by Ana Cláudia Duarte Cardoso, and José Júlio Ferreira Lima, 145–172. Rio de Janeiro: Observatório das metrópoles: Letra. capital. <http://livroaberto.ufpa.br/jspui/handle/prefix/372>.
- Meier, Patrick. 2015. *Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response*, 1st ed. Oxfordshire: Routledge. <https://doi.org/10.1201/b18023>.
- Pegado, Rosielle and Claudio Blanco, Jackson Roehrig, Carla Carocha, and Francisco Costa. 2014. *Risco de Cheia e Vulnerabilidade: Uma Abordagem às Inundações Urbanas de Belém/Pará/Brasil*. Revista Territorium. 21. 71. https://doi.org/10.14195/1647-7723_21_6.
- Penteado, Antonio Rocha. 1968. *Belém: (estudo de geografia urbana)*. Belém: UFPA. 2 v. (Coleção amazônica. Série José Veríssimo). Available at <http://livroaberto.ufpa.br/jspui/handle/prefix/43>.
- Rodrigues, Edmilson Brito. 1996. *Aventura Urbana: Urbanização, Trabalho e Meio Ambiente em Belém*. Belém: Editora.
- Silva, Rogério Renato, Max Gasparini, Elis Alquezar, Paola Gongra, and Antonio Ribeiro. 2017. *Avaliação para Negócios de Impacto Social: Guia Prático*. São Paulo: Atermisia.
- Zuckerman, Ethan. 2014. "New Media, New Civics?" *Policy & Internet* 6 (2): 151–168. <https://doi.org/10.1002/1944-2866.POI360>.

About the Authors

Acilon Himercício Baptista Cavalcante is an architect and interaction designer with a Master's degree in architecture and urbanism and a Master's in Arts. He is a doctoral candidate in Digital Media at the University of Porto, Portugal. He is an assistant professor at the Federal University of Pará (UFPA), Brazil.

Ana Cláudia Duarte Cardoso is an architect with a Master's degree in Urban Planning and a Ph.D. in Architecture. She is a Full Professor at the Federal University of Pará (UFPA), Brazil, and the Graduate Program in Architecture and Urbanism at the same institution. She participated in social movements, public administration, and multidisciplinary research.

13. Speculative Data Infrastructures: Prototyping a Public Database on Corporate Tax Avoidance

Jonathan W. Y. Gray

Abstract

This chapter examines and reflects on three ways of workshopping ‘data in the making’ – redesigning, prototyping, and interfacing – drawing on activities leading towards a collaborative public database on the economic activities and tax contributions of multinational corporations. It examines how prototyping data infrastructures may serve as a method to engage with organizations, groups, and communities who are concerned with or affected by an issue, in order to materialize problems and to support learning which may go on to inform advocacy, policy, and reporting activities. It draws on ‘engaged research-led teaching’ activities with King’s College London, the Public Data Lab, and the Tax Justice Network to consider formats for critically engaging with data as a medium for issue articulation.

Keywords: Data studies; Participatory design; Engaged teaching; Tax justice

Introduction

How much tax do multinational corporations really pay? How much are they avoiding? What difference could it make if national and international rules were strengthened in order to tackle multinational tax avoidance? Since the turn of the millennium, tax justice advocates have argued for Country-by-Country Reporting (CBCR) standards that would disclose how corporations

avoid tax, including through profit-shifting and other mechanisms (Kohonen and Mestrum 2008; Leaman and Waris 2013; Tax Justice Network 2015; Seabrooke and Wigan 2015b). Although earlier proposals were initially widely critiqued and rejected by many accountancy professionals, CBCR has subsequently made it into law and policymaking around the world, including at transnational bodies such as the OECD and the European Union (EU).¹ Most multinational firms do not yet have to publish CBCR data, but some of them do – particularly those operating in sectors where there are CBCR rules, such as extractive industry companies in the United States and financial institutions with operations in the EU.²

This chapter explores some of the things we have learned from an overlapping series of projects and activities exploring and critically engaging with CBCR data over a five-year period (2016–2021). It focuses on a series of workshops with CBCR data undertaken with members of the Tax Justice Network and the Open Data for Tax Justice initiative, as well as students and researchers at King’s College London and the Public Data Lab.³ Social and humanities researchers are used to ‘workshop’ texts. But what kinds of formats, we should ask, might be used to workshop data? Furthermore, how might one workshop data when there is not already a dataset or data collection to hand around, but where data is scattered, partial, not-yet-completed, or “in the making” (Latour 1988)? Thinking along with recent work on formats such as the “data sprint” and the “technocultural workshop” to support encounters with digital data, methods, and devices (Munk, Meunier, and Venturini 2019; Venturini, Munk, and Meunier 2018; Berry et al. 2015; Coté and Pybus 2016), the following sections explore three ways of workshoping data-in-the-making: redesigning, prototyping, and interfacing.

In alignment with work on data feminism (D’Ignazio and Klein 2020), situated data analysis (Rettberg 2020), data hermeneutics (Acker 2015; Gerbaudo 2016; Poirier 2021), and participatory data design (Jensen et al. 2021), these

1 See, for example, OECD BEPS Action 13 <https://www.oecd.org/tax/beps/beps-actions/action13/> and EU CBCR rules pertaining to multinationals, extractive and logging industries: https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/public-country-country-reporting_en.

2 This includes, for example, Section 1504 of the Dodd Frank Act in the United States, and the EU Capital Requirements Directive IV (CRD IV).

3 Thanks to Liliana Bounegru, Alex Cobham, Javier Garcia-Bernardo, Cristina Fernandez, Franki Hackett, Petr Janský, Danny Lämmerhirt, Anna Powell-Smith, Stephen Abbott Pugh, our data journalism students at the Department of Digital Humanities, King’s College London, colleagues at the Public Data Lab, and all who took time to advise, support and be involved in these workshops in various ways. For more information, see <https://publicdatalab.org/projects/corporate-tax-data/>.

approaches are intended to support critical engagements with the question of what role data may play in relation to societal issues – engagements that are grounded in social and cultural research as well as in collaborations with activist, civil society, and community groups.

Who and What Is Data For? Redesigning Data Infrastructures

We started with a series of smaller workshops with researchers and activists on the prospects of a public database dedicated to the economic activities and tax contributions of multinationals. These workshops took place against a background of research on the politics of public data that emphasizes the ‘opening up’ of existing institutional data, towards experiments in citizen and civil society data gathering and engagements (Gray 2016; Gray, Lämmerhirt, and Bounegru 2016; Gray 2018) as well as what we have called “data infrastructure literacy” which is “not just competencies in reading and working with datasets but also the ability to account for, intervene around and participate in the wider socio-technical infrastructures through which data is created, stored and analysed” (Gray, Gerlitz, and Bounegru 2018, 1).

Rather than starting with datasets and exploring what can be done with them – the premise of many hackdays, hackathons, challenges, and fellowships – we started with an approach that sought to prioritize *communities, issues* and *questions*. The aspiration was to be community-centered, issue-centered and question-centered rather than data-centered. Engaging with civil society groups, nonprofits, and researchers, we sought to elicit pressing questions as well as the contexts in which these questions arose. An overview of the lines of inquiries and approaches from these workshops is included in Table 13.1.

Community snowballing with workshops, conversations, and online materials were used to gather a network of people who were interested in using public data on corporate tax avoidance in their work, which later became the Open Data for Tax Justice initiative.⁴ With this broader group we could gather questions that related to their work: Which companies are avoiding tax in my country? How much profit is declared by mining and petroleum companies in countries where resources are extracted? In other countries? In tax havens? Workshop participants collectively drafted ‘user stories’ as a way to understand more about the contexts

4 See <https://datafortaxjustice.net/>.

Table 13.1 Workshopping approaches for mapping communities, questions, circumstances, data proposals, and existing data

Lines of inquiry	Workshopping approaches
Which communities are concerned with and affected by the issue?	Community snowballing with workshops, interviews, online materials.
What questions do these communities have about the issue?	Developing lists of questions across workshops and sharing to elicit further input.
How, for whom, and in which circumstances might data help to address these questions?	Collective user story exercises to elicit more situated account of contexts in which questions arise and where data could help.
What data is proposed in order to make these questions addressable?	Mapping and comparing proposed data models and data standards, further developing proposals with materials above and soliciting for further input.
How does existing data compare with desired data?	Mapping existing data, the contexts in which it is produced, and assessing this against proposals above to surface gaps and differences.

in which these questions mattered. For example: “As a [X] I need/want/expect to [Y] so that [Z].” While user stories can be used to abstract and fix interests for the purposes of making software or products, for our data workshops they were also a way to situate and pluralize concerns – and to be alive to many kinds of interests, issues, and settings that brought communities to care.⁵

These communities, questions, and circumstances then served as a starting point from which to assess data proposals as well as existing data – and in the process to surface gaps, differences, shortcomings, misalignments, unaddressed concerns, and unanswerable questions (figure 13.1). This workshopping process also drew attention to the varied contexts of data fields and what they depended upon (e.g., thresholds, definitions), as well as whose concerns were better represented and whose were missing.

The results of these initial workshops and mapping activities were published in a report with the Tax Justice Network (Cobham, Gray, and Murphy 2017). Collaborators at the Tax Justice Network have said that learning from this collaboration has shaped their advocacy activities and thinking around the role of public data in addressing tax justice, including key contributions

5 This mode of working with user stories has been written about in collaborative design research drawing on science and technology studies (Poderi et al. 2020).

Activity

Civil society proposal	OECD CBCR	CRD IV	Dodd Frank	Canada	EITI	EJ
Third party sales	✓	✗	✗	✗	✗	✗
Turnover	✓	✓	✗	✗	✗	✗
Number of employees FTE	✓	✓	✗	✗	✗	✗
Total employee pay	✗	✗	✗	✗	✗	✗
Assets?	✗	✗	✗	✗	✗	✗

Figure 13.1. Table showing which data elements were missing from which data proposals. Source: Open Data for Tax Justice.

to the Global Reporting Initiative’s tax standard. The approach was also used to workshop data with other communities, such as public housing campaigners.

How Could a Database Work? Testing Transparency and Prototyping Collaborative Data Infrastructures

Further workshops were dedicated to exploring how a corporate tax database could work in practice. These workshops were undertaken in the context of “engaged research-led teaching” activities, involving researchers, students, and civil society actors such as activists and journalists as co-inquirers in a way which cared for their various perspectives, needs, and concerns (Gray et al. 2022). Several workshops took place as part of a data journalism module at King’s College London, for which the main text was *The Data Journalism Handbook: Towards a Critical Data Practice*, which sought to encourage “a relational perspective on data journalism as a kind of curatorial craft, assembling and working with diverse materials, communities and infrastructures to generate different ways of knowing, narrating and seeing the world at different scales and temporalities” (Bounegru and Gray 2021). The book highlighted how journalists not only used existing datasets but were increasingly assembling their own data.

One of the options for student group projects was to work together to prototype a collaborative data infrastructure for learning about corporate tax avoidance. The idea was to take data that was available in principle through

a patchwork of different rules (e.g., EU Capital Requirements Directive IV) but often scattered in corporate websites and PDFs, and to explore how different actors could work together to assemble and use it while also being able to trace where it had come from and assess the many issues associated with it. Several research groups had compiled subsets of this data (Janský 2020), but it was felt that these efforts had been somewhat ad hoc: further work was needed to improve the coordination and documentation of what was there and what was missing, beyond and across projects.

In 2016 we had a daylong workshop with members of the Tax Justice Network and data journalists who were present for the duration. By 2021, being mindful of the limited time of our collaborators, we had packages of materials and documentation that researchers and students could work with independently across half-day workshops, receiving input and feedback at key moments. These events involved multiple smaller groups working on and coordinating around different aspects of assembling data – making lists of companies who had to comply with rules, gathering and publicly archiving copies of PDF reports, making guides to finding CBCR data, assessing and trying out PDF transcription tools, extracting data tables from PDF reports, assessing and trying out database tools which could be used for a collaborative database, assessing and trying out tools and interfaces to support distributed collaboration around the data such as transcription, micro-tasking, and forums (figure 13.2).

In the context of a data journalism course grounded in critical data studies, science and technology studies, and associated fields, it was important for our students' learning that these tasks were not just instrumental actions but also opportunities for "critical proximity" (Birkbak, Petersen, and Elgaard Jensen 2015); here we drew on class readings on information infrastructures, classification and standards, histories and sociologies of quantification, "statactivism" and data activism (e.g., Bowker and Star 2000; Espeland and Stevens 2008; Rottenburg, Merry, Park, and Mugler 2015; Lampland and Star 2009; Bruno, Jany-Catrice, and Touchelay 2016; Desrosières 2002; Merry 2016; Bruno, Didier, and Vitale 2014; Milan and van der Velden 2016).⁶ The students' challenge was to look at how these readings might become salient in engagements with data, including in advocacy and policy concerned with the making of data. These workshops and activities resulted in packages of materials with prototypes, documentation, and options for next steps for a collaborative data infrastructure. The workshops emphasized not

6 For additional example readings see https://www.zotero.org/groups/sociology_of_quantification and https://www.zotero.org/groups/data_journalism_research.

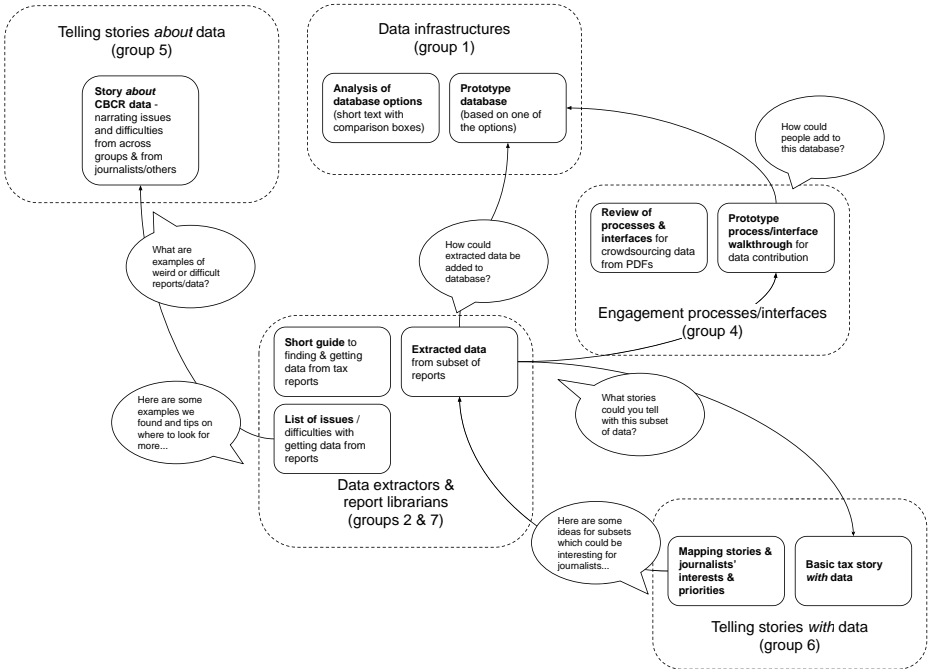


Figure 13.2. Group projects on assembling tax data at King’s College London. Source: Data journalism class of 2021, King’s College London.

only technical and statistical literacy but also relational perspectives on data infrastructures and the social lives of data. Included here were data issues and data frictions (Edwards et al. 2011) that had been identified throughout the workshops, such as broken links, missing documents, missing fields, inconsistent units, divergent accounting practices, and rejected FOI requests.⁷

By having multiple projects take different prototyping approaches, participants were able to learn from one another and identify considerations for future work grounded in various kinds of hands-on engagements – such as querying, scraping, transcribing, archiving and interpreting reports. Workshopping data through prototyping involved the testing and empirical re-specification of transparency measures and their effects – often surfacing scattered, heterogeneous materials rather than revealing a clearer, bigger

7 A student FOI request aiming to establish which corporations fell under CBCR rules led to a government response that multinationals were considered ‘persons’ and therefore this could not be disclosed, which an activist collaborator described as a “shocking” rolling back of previous transparency commitments.

picture on corporate tax avoidance.⁸ Tax Justice Network collaborators commented that these workshops represented the “single biggest contribution” to work in this area since the report mentioned in the previous section.

How Could It Be Meaningful? Outlining Interfaces and Data Stories

What could be done with CBCR data? A third format for workshoping data involved the outlining of experimental interfaces and data storytelling approaches. Researchers have drawn attention to the role of formats such as benchmarks and indices in tax justice advocacy (Seabrooke and Wigan 2015a). What kinds of formats for displaying, exploring, and making sense of data could be appropriate given what had been learned about its problems, partiality, provisionality, and incompleteness?

These data workshops took their cue from “challenges for critical data practice” in the *Data Journalism Handbook*, including the following questions:

How can data journalism projects tell stories both with and about data including the various actors, processes, institutions, infrastructures and forms of knowledge through which data is made? [...] How can data journalism projects tell stories about big issues at scale (e.g., climate change, inequality, multinational taxation, migration) while also affirming the provisionality and acknowledging the models, assumptions and uncertainty involved in the production of numbers? [...] How can data journalism projects cultivate their own ways of making things intelligible, meaningful and relatable through data, without simply uncritically advancing the ways of knowing “baked into” data from dominant institutions, infrastructures and practices? (Bounegru and Gray 2021)

These workshops drew on projects and practices such as Mona Chalabi’s approach to data sketching (Chalabi and Gray 2021), a Public Data Lab collaboration exploring the visual representation of uncertainty in offshore finance,⁹ and design practices to mock-up interfaces for not-yet-existent data infrastructures.

8 Commensurate with what other researchers and investigators have found with transparency initiatives in other areas, e.g., <https://blog.okfn.org/2011/03/08/a-kafkaesque-data-trail-the-hunt-for-europes-hidden-billions/>.

9 See <https://publicdatalab.org/projects/fog-of-finance/> and <https://offshoreatlas.publicdatalab.org/>.

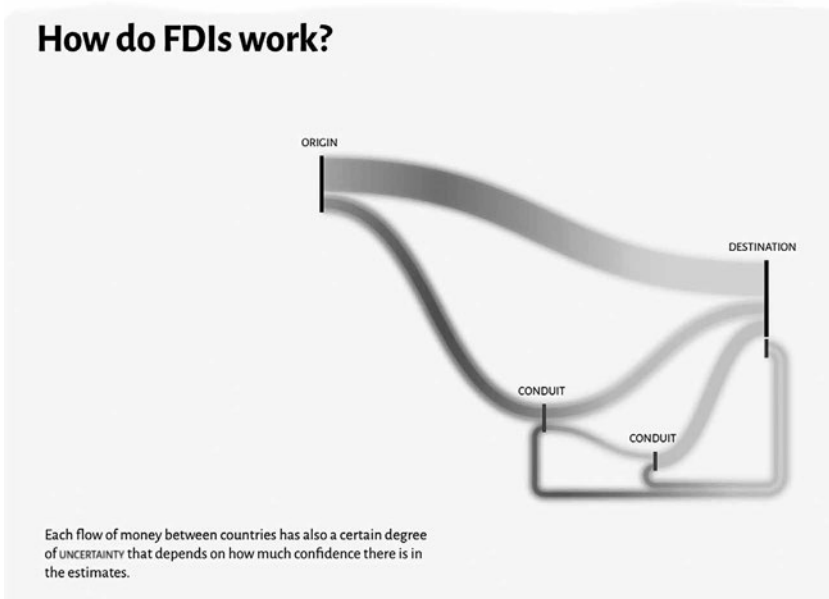


Figure 13.3. Exploring the visualization of uncertainty with the “Fog of Finance” project. Source: Public Data Lab.

In these workshops, participants outlined storytelling approaches and possible interfaces, visuals, and other materials which could be used to make sense of corporate tax data. This activity often started by longlisting stories where there were big questions (such as those in the “What Do They Pay?” report) that they discovered could not be answered. Some groups explored the visual display of possible indicators of tax avoidance – such as data visualizations of ratios of profits to employees for each jurisdiction of each corporate group to identify outliers (corporate entities with high ratios of profits to employees may indicate avoidance). These projects proceeded by identifying meaningful and doable subsets to tell stories about – such as top corporate entities in particular sector by size. Other groups sketched interfaces for making games with data or inviting involvement in gathering it.

Along with such approaches to telling stories and making things ‘with data’, other groups looked at telling stories ‘about data.’ Inspired by Helen Verran’s chapter in the *Data Journalism Handbook* “Narrating a Number and Staying With the Trouble of Value” (Verran 2021) and by readings on the politics of data and quantification, workshopping data became not just an opportunity to regard data as material for stories about the entities within it (corporations, transactions, jurisdictions) but as an invitation to take

data as an entry point for telling stories about the codification of concern, expert knowledge cultures, accounting, the emergence of the corporate form, identifiers and thresholds, problematization and answerability, data imaginaries, data politics, and data-in-the-making. Rather than taking data as background work for the making of portable facts, numbers, claims, and evidence, such stories could re-animate data as a site of struggle and contestation around the making and ordering of economic life and collective futures.

Conclusion

The three forms of workshopping data-in-the-making presented above suggest ways of organizing collaborations with data beyond a focus on capacities that may be conventionally desired or expected for its effective manipulation (e.g., data science, programming). They indicate other kinds of collective learning – including situating how and for whom data infrastructures matter, surfacing frictions and considerations for collaborative infrastructures, and sketching outlines for alternative interfaces and data storytelling approaches.

In the case of corporate tax avoidance, they helped to foment encounters between researchers, teachers, students, and activists around the role that data may play in relation to materializing and addressing corporate tax avoidance, surfacing who and what is missing from proposals as well as the messiness of data and accounting practices, and unveiling alternative approaches for telling stories – provisional, partial, and incomplete, but nevertheless revealing. These kinds of data workshops can be a generative format for collective learning and interpretation, shaping expectations and orienting activities (de Mourat, Ricci, and Latour 2020), not just a means for producing outcomes or outputs which are known in advance.

Workshop formats have their own politics, and their conventions can embody problematic defaults which shape, silence, and order involvement – foregrounding and marginalizing, distributing work unfairly, making visible certain kinds of voices, concerns, and experiences while making other kinds invisible (Salesses 2021; Pierre et al. 2021). If data workshop formats are to be inclusive, meaningful, and equitable, they should be considered, documented, and cared for, in order to assess whether they are working for those involved and affected.

References

- Acker, Amelia. 2015. "Toward a Hermeneutics of Data." *IEEE Annals of the History of Computing* 37 (3): 70–75. <https://doi.org/10.1109/MAHC.2015.68>.
- Berry, David M., Erik Borra, Anne Helmond, Jean-Christophe Plantin, and Jill Walker Rettberg. 2015. "The Data Sprint Approach: Exploring the Field of Digital Humanities through Amazon's Application Programming Interface." *Digital Humanities Quarterly* 9 (3).
- Birkbak, Andreas, Morten Krogh Petersen, and Torben Elgaard Jensen. 2015. "Critical Proximity as a Methodological Move in Techno-Anthropology." *Techné: Research in Philosophy and Technology* 19 (2): 266–290. <https://doi.org/10.5840/techné201591138>.
- Bounegru, Liliana, and Jonathan Gray, eds. 2021. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press. <https://www.aup.nl/en/book/9789462989511/the-data-journalism-handbook>.
- Bowker, Geoffrey C., and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bruno, Isabelle, Emmanuel Didier, and Tommaso Vitale. 2014. "Statactivism: Forms of Action Between Disclosure and Affirmation." *Partecipazione e Conflitto – Participation and Conflict* 7 (2): 198–220.
- Bruno, Isabelle, Florence Jany-Catrice, and Béatrice Touchelay (eds). 2016. *The Social Sciences of Quantification: From Politics of Large Numbers to Target-Driven Policies*. New York, NY: Springer.
- Chalabi, Mona, and Jonathan Gray. 2021. "Sketching with Data." In *The Data Journalism Handbook: Towards a Critical Data Practice*, edited by Liliana Bounegru and Jonathan Gray, 174–181. Amsterdam: Amsterdam University Press. <https://www.aup.nl/en/book/9789462989511/the-data-journalism-handbook>.
- Cobham, Alex, Jonathan Gray, and Richard Murphy. 2017. "What Do They Pay? Towards a Public Database to Account for the Economic Activities and Tax Contributions of Multinational Corporations." January 2017. CIYPERC Working Paper Series. London: City University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3049857.
- Coté, Mark, and Jennifer Pybus. 2016. "Simondon on Datafication. A Techno-Cultural Method." *Digital Culture & Society* 2 (2): 75–92. <https://doi.org/10.14361/dcs-2016-0206>.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Desrosières, Alain. 2002. *The Politics of Large Numbers: A History of Statistical Reasoning*, translated by Camille Naish. Cambridge, MA: Harvard University Press.

- Edwards, Paul N., Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman. 2011. "Science Friction: Data, Metadata, and Collaboration." *Social Studies of Science* 41 (5): 667–90. <https://doi.org/10.1177/0306312711413314>.
- Espeland, Wendy Nelson, and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology / Archives Européennes de Sociologie* 49 (3): 401–436.
- Gerbaudo, Paolo. 2016. "From Data Analytics to Data Hermeneutics. Online Political Discussions, Digital Methods and the Continuing Relevance of Interpretive Approaches." *Digital Culture & Society* 2 (2): 95–112. <https://doi.org/10.14361/dcs-2016-0207>.
- Gray, Jonathan. 2016. "Datafication and Democracy: Recalibrating Digital Information Systems to Address Societal Interests." *Juncture* 23 (3): 197–201. <https://doi.org/10.1111/newe.12013>.
- Gray, Jonathan. 2018. "Three Aspects of Data Worlds." *Krisis: Journal for Contemporary Philosophy*, no. 1. <https://archive.krisis.eu/three-aspects-of-data-worlds/>.
- Gray, Jonathan, Liliana Bounegru, Richard Rogers, Tommaso Venturini, Donato Ricci, Axel Meunier, Sabine Niederer, Natalia Sánchez Querubín, Marc Tuters, Lucy Kimbell, and Anders Kristian Munk. 2022. "Composing Collective Inquiry with Digital Methods and Data." *Digital Culture & Education (DCE)* 14 (3): 55–86.
- Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. "Data Infrastructure Literacy." *Big Data & Society* 5 (2): 1–13. <https://doi.org/10.1177/2053951718786316>.
- Gray, Jonathan, Danny Lämmerhirt, and Liliana Bounegru. 2016. "Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection?" CIVICUS and Open Knowledge International. <http://papers.ssrn.com/abstract=2742871>.
- Janský, Petr. 2020. "European Banks and Tax Havens: Evidence from Country-by-Country Reporting." *Applied Economics* 52 (54): 5967–5985. <https://doi.org/10.1080/00036846.2020.1781773>.
- Jensen, Torben Elgaard, Andreas Birkbak, Anders Koed Madsen, and Anders Kristian Munk. 2021. "Participatory Data Design: Acting in a Digital World." In *Making & Doing: Activating STS through Knowledge Expression and Travel*, edited by Gary Lee Downey and Teun Zuiderent-Jerak, 117–136. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/11310.001.0001>.
- Kohonen, Matti, and Francine Mestrum, eds. 2008. *Tax Justice: Putting Global Inequality on the Agenda*. London: Pluto Press.
- Lampland, Martha, and Star, Susan Leigh. 2009. *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.
- Latour, Bruno. 1988. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.

- Leaman, Jeremy, and Attiya Waris, eds. 2013. *Tax Justice and the Political Economy of Global Capitalism, 1945 to the Present*. New York, NY: Berghahn Books.
- Merry, Sally Engle. 2016. *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago, IL: University of Chicago Press.
- Milan, Stefania, and Lonneke van der Velden. 2016. "The Alternative Epistemologies of Data Activism." *Digital Culture & Society* 2 (2): 57–74.
- Mourat, Robin de, Donato Ricci, and Bruno Latour. 2020. "How Does a Format Make a Public?" In *Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access*, edited by Martin Paul Eve and Jonathan Gray, 103–112. Cambridge, MA: MIT Press.
- Munk, Anders Kristian, Axel Meunier, and Tommaso Venturini. 2019. "Data Sprints: A Collaborative Format in Digital Controversy Mapping." In *DigitalSTS: A Field Guide for Science and Technology Studies*, edited by Janet Vertesi and David Ribes, 472–496. Princeton, NJ: Princeton University Press.
- Pierre, Jennifer, Roderic Crooks, Morgan Currie, Britt Paris, and Irene Pasquetto. 2021. "Getting Ourselves Together: Data-Centered Participatory Design Research & Epistemic Burden." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–11. CHI '21. New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445103>.
- Poderi, Giacomo, Hanna Hasselqvist, Andrea Capaccioli, Cristian Bogdan, and Vincenzo D'Andrea. 2020. "Matters of Concerns and User Stories: Ontological and Methodological Considerations for Collaborative Design Processes." *CoDesign* 16 (3): 220–232. <https://doi.org/10.1080/15710882.2018.1557694>.
- Poirier, Lindsay. 2021. "Reading Datasets: Strategies for Interpreting the Politics of Data Signification." *Big Data & Society* 8 (2): 20539517211029320. <https://doi.org/10.1177/20539517211029322>.
- Retberg, Jill Walker. 2020. "Situated Data Analysis: A New Method for Analysing Encoded Power Relationships in Social Media Platforms and Apps." *Humanities and Social Sciences Communications* 7 (1): 1–13. <https://doi.org/10.1057/s41599-020-0495-3>.
- Rottenburg, Richard, Sally E. Merry, Sung-Joon Park, and Johanna Mugler, eds. 2015. *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge: Cambridge University Press.
- Salesses, Matthew. 2021. *Craft in the Real World: Rethinking Fiction Writing and Workshopping*. New York, NY: Catapult.
- Seabrooke, Leonard, and Duncan Wigan. 2015a. "How Activists Use Benchmarks: Reformist and Revolutionary Benchmarks for Global Economic Justice." *Review of International Studies* 41 (5): 887–904. <https://doi.org/10.1017/S0260210515000376>.
- Seabrooke, Leonard, and Duncan Wigan. 2015b. "Powering Ideas through Expertise: Professionals in Global Tax Battles." *Journal of European Public Policy* 23 (3): 357–374. <https://doi.org/10.1080/13501763.2015.1115536>.

- Tax Justice Network. 2015. *The Greatest Invention: Tax and the Campaign for a Just Society*. Margate: Commonwealth Publishing.
- Venturini, Tommaso, Anders Munk, and Axel Meunier. 2018. "Data-Sprinting: A Public Approach to Digital Research." In *Handbook of Interdisciplinary Research Methods*, edited by Celia Lury, Rachel Fensham, Alexandra Heller-Nicholas, Sybille Lammes, Angela Last, Mike Michael, and Emma Uprichard. London: Routledge.
- Verran, Helen. 2021. "Narrating a Number and Staying With the Trouble of Value." In *The Data Journalism Handbook: Towards A Critical Data Practice*, edited by Liliana Bounegru and Jonathan Gray, 55–64. Amsterdam: Amsterdam University Press. <https://doi.org/10.2307/j.ctv1qr6smr.11>.

About the Author

Jonathan W. Y. Gray is Senior Lecturer in Critical Infrastructure Studies at the Department of Digital Humanities, King's College London. He is also Cofounder of the Public Data Lab; and Research Associate at the Digital Methods Initiative (University of Amsterdam) and the médialab (Sciences Po, Paris).

14. The DataWorkplace: Collaborative Learning about Datafication in Local Government

Krista Ettliger, Mirko Tobias Schäfer, Albert Meijer, and Martiene Branderhorst

Abstract

The DataWorkplace is a transdisciplinary research collaboration between Utrecht University and a group of local and regional government organizations in the Netherlands. Its goal is to investigate and to support the digital transformation of public management in local and regional government organizations. The transdisciplinary quality of the DataWorkplace manifests in its hands-on, participation-intensive style of collaborative research; academic researchers and practitioners from local and regional government organizations collaborate on an equal basis while exploring pressing research questions together. This chapter presents outcomes of this ongoing cooperation and discusses the benefits and challenges of engaging in transdisciplinary work for academics, practitioners, networked collaborators, and society at large.

Keywords: Digital transformation; Public management; Good governance; Transdisciplinary research

Introduction

Datafication presents a complex challenge for government organizations. On the one hand, the increasing availability of data (Kitchin 2014) offers opportunities to government organizations: it can help them improve

their services, and citizens can be empowered and their quality of life enhanced (Kim and Chung 2014; Redden 2018; Meijer 2018; Micheli et al. 2020). On the other hand, datafication entails operational risks for government organizations, including the inability to handle and process such data (Kim and Chung 2014; Maciejewski 2017), as well as ethical risks in terms of transparency, privacy, and fairness (Maciejewski 2017; Redden 2018). Thus, the challenge for government organizations lies in facing that not only is the technology new and requires technological capabilities to work with it, but also that increasing datafication raises a host of other questions about the organizations themselves, including their structure and way of operating (Branderhorst 2020; Meijer 2018), their organizational culture (Redden 2018), their legitimacy (Klievink et al. 2017; Grimmelikhuijsen and Meijer 2022), and their ethics (Redden 2018; Siffels et al. 2022).

The datafication of local and regional government is situated in the midst of a profound transformation of public management in the Netherlands. This change is being driven not only by technological development and the digital transition but also by policy, such as the central government's delegation of tasks to municipalities, and by major administrative objectives, such as the energy transition, climate action, housing targets, the mobility agenda, etc. The use of data and algorithms are at the center of innovations in public management and in responses to contemporary challenges. Instead of studying these developments from afar, the DataWorkplace investigates, in an exploratory fashion, how these challenges manifest themselves within organizations, and it develops a conceptualization of good digital governance while initiating applicable solutions.

The Dutch Research Council rejected the initial proposal for the DataWorkplace on the grounds that its ability to connect to stakeholders was claimed to be limited. Those stakeholders, however, decided to finance the project themselves. The DataWorkplace was created as a collaboration between the Utrecht School of Governance and the Data School, both at Utrecht University, along with four local and regional government partners: the municipality of Gouda, the municipality of Almere, the municipality of Woerden, and the province of South Holland. These organizations participated in the collaboration from the beginning of 2019 until the end of 2020. In January 2021 a second round of the DataWorkplace began, keeping three of the existing government partners – the municipalities of Gouda and of Almere, as well as the province of South Holland – and gaining three new government partners: the municipalities of Amersfoort and of Zuidplas, along with the province of Utrecht. The third and current round involves

the provinces of Utrecht, North Brabant, and South Holland, as well as the municipalities of Amersfoort and Utrecht.¹

In this chapter's case study, we explore the DataWorkplace project, which focuses on transdisciplinary and collaborative research to address the complexities introduced by datafication in local and regional governments in the Netherlands. Our discussion centers on both the benefits and the challenges encountered in this research approach.

Collaborative Research

Organizing research projects at the DataWorkplace differs from what transpires with traditional academic research projects right from the start. Here, a steering committee, consisting of the academic PIs and high-level employees from the partner organizations, identifies possible research questions in a collaborative design session. The researchers inform the process with theoretical grounding, connecting the issues brought forward by the extra-university partners to research agendas in governance studies and critical data studies (Van Es and Schäfer 2017; Schäfer, Van Es, and Muis 2023; Meijer, Ingrams, and Zouridis 2022). Contemporary academic discussions are closely aligned with the immediate demands and requirements of societal sectors, particularly public management organizations grappling with datafication challenges. This alignment also enables us to dynamically adjust our research focus in response to organizational shifts or technological advancements (e.g., Ruijter, Dingelstad, and Meijer 2023).

Our steering committee meetings played a crucial role in uncovering varied motivations for participating in the collaboration. Over time, we identified two primary agendas. First, there is an urgent need for practical solutions in tackling pressing challenges arising from technological advancements and in effectively implementing new technologies. Second, a more profound interest has emerged towards understanding the theory of change to promote effective governance practices in digital public administration and to deepen our understanding of how datafication impacts government operations.

The research activities are collaboratively undertaken by academics and civil servants alike. Students from bachelor and master's levels work together with the senior researchers and the government partners. Students are

¹ Since then, the Dutch Research Council has featured the DataWorkplace as an example of alternatively funded research projects. We wear this acknowledgement as a badge of honour.

recruited from the fields of governance studies, media studies, and applied ethics. Once the topics are determined, researchers and mid-level contacts within the organizations co-design the research. During the research process, iterative meetings with stakeholders explore opportunities to develop applications that can help solve some of the issues under investigation or that were revealed during the investigation. This means that instead of delivering the research results as a report, researchers and practitioners translate them into concrete instruments that employees in the organization can use. At the outset, this practicality was one of the unique characteristics of the DataWorkplace – instead of only transmitting reports that just create more paperwork for practitioners, it delivers tools to help these organizations embed their data practices within the organization.

Another unique feature of the DataWorkplace is the network element, connecting the participating organizations with each other and connecting organizations with researchers. Collaboration with researchers is an important aspect in the development of new capabilities within local government organizations (Rathenau 2020). However, it can be difficult for smaller municipalities to finance such research efforts (*idem*). For researchers it is also not always easy to create effective knowledge transfer, affecting whether their findings are translated into applicable solutions. Here the DataWorkplace offers a solution to these difficulties. By being structured as a network, organizations can pool their resources and collaborate with university partners. Additionally, because the research is shared between the organizations, they also benefit from learning from and about the organizations within the network (Ettlinger 2021).

Finally, the DataWorkplace is directly connected to academic practices such as international conferences in public administration and media studies, as well as international network collaborations with other researchers and publications in national and international journals. These sorts of connections are important for the continuous confirmation of research quality and for linking our contextual analyses to broader academic debates.

Knowledge Transfer and the Application of Research Findings

Because the DataWorkplace strives to facilitate effective knowledge transfer and the immediate application of research findings, problem-solving is central to the collaborative effort it undertakes. During the first round of collaboration, the DataWorkplace developed five instruments with the participating organizations. These were: a process for setting up a data

team called *Data Team Start*; a checklist consisting of aspects to regard when sharing data (*Samen Data Delen*); and a test for measuring the awareness of data ethics in an organization (*Data Ethics Awareness Test*).² In addition, there were guidelines and a workshop on the meaningful use of dashboards.³ Motivated by the challenges of remote labor arising during the pandemic, guidelines for meaningful digital and hybrid collaboration were developed.⁴ The guidelines for both remote collaboration and for using dashboards responded to current and pressing problems in the organizations and were distributed so as to inform and change practices. They also show how stakeholder needs directly inform the research process. In the aforementioned steering committee, practitioners directly address current challenges and emerging urgencies, enabling the allocation of resources to investigate these issues. This approach fosters dynamic research processes and generates research questions that might not be evident without input from external partners.

Over the past five years, the researchers at the Data Workplace have investigated a broad range of topics that include data team collaboration; data visualization for better policy; responsible data practices by municipalities; digital twins and related governance challenges; the government as platform; and citizen participation through data projects. Recently the topic of using large language models in public management organizations has been added to the agenda. Two dynamics evolve in parallel here: short-term projects revolving around a case example or a specific technology, and a long-term perspective on good governance for digital society. The short-term projects are well suited for the master theses of student researchers and support the partner organization in developing applicable solutions to these challenges. The case examples as well as the application inform the long-term perspective on good governance practices. This leads to four different forms of research output: 1) graduation projects by research master's students; 2) information and applicable solutions for internal use by the partner organization; 3) articles in professional magazines in the Dutch public administration sector; and 4) academic publications in peer-reviewed journals.

2 For an overview, see the product page of the DataWorkplace: <https://datawerkplaats.org/producten-en-tools/>.

3 DataWerkplaats: Dashboards met waarde voor de hele organisatie [Dashboards with value for the entire organisation], <https://datawerkplaats.org/producten-en-tools/dataweergave/dashboards-met-waarde-voor-de-hele-organisatie-concept/>.

4 <https://datawerkplaats.org/producten-en-tools/archief/zinvol-digitaal-en-hybride-samenwerken/>.

The first rounds of the DataWorkplace were strictly tool-focused, but the collective learning process caused a gradual shift to take place. We noticed that the tools, however much they may have helped develop a range of important insights, were not necessarily being used well by organizations. Tool design, as a form of transdisciplinary learning embedded in a specific context, was important in strengthening the organizational understanding of socio-technical processes of change; nevertheless, the tools did not always travel well to other contexts. We therefore started putting less emphasis on the tools as such and focused more on generating specific knowledge for a particular situation given its context, and we would then find other formats – presentations, posters, papers – to convey this knowledge to other government partners.

Some tools, however, did in fact travel well to other contexts – these were the Data Ethics Decision Aid (DEDA) and the Code Digital Government (CODIO).⁵ DEDA is a process for dialogic deliberation on the design of a data or AI project (Franzke, Muis, and Schäfer 2021), and CODIO is a workshop helping civil servants and policymakers to map good governance practice for digital solutions in public management (Meijer and Ruijter 2021). Both are accompanied by a particular type of workshop with staff trained at Utrecht University, designed to meet the demands of training and application in the field. The public management sector in the Netherlands makes extensive use of both DEDA and CODIO, which seem not only better supported and more thoroughly developed but also more optimally aligned with the focus on organizational change. Having moved away from solely addressing particular local problems, the consortium of the DataWorkplace now concentrates on issues common to several of the participating organizations.

An important format for knowledge transfer in the DataWorkplace is the so-called knowledge conference, which takes place twice a year. Hosted by one of the partner organizations, this conference is a venue where researchers give presentations about the progress of their work. In addition, these meetings are used for collective learning about the challenges that the digital transition poses to public management. Usually, the audience is not limited to direct DataWorkplace participants and members of their organizations. Civil servants from other municipalities and provinces attend these knowledge conferences to learn about research projects and to connect

5 See Utrecht University / Data School: Data Ethics Decision Aid (DEDA), online: <https://deda.dataschool.nl/en/> and Utrecht University / DataWorkplace: Code Goed Digitaal Openbaar Bestuur (CODIO), online: <https://datawerkplaats.org/producten-en-tools/reflectie-evaluatie-verantwoorde-digitalisering/goed-digitaal-bestuur/>.

with the learning community. These opportunities point to an important motivation for joining the DataWorkplace. Although the research projects and their results provide one motivation to participate in the DataWorkplace, the community that has evolved around it has become an almost equally important attractor.

Observed Benefits and Challenges of Transdisciplinary Research

The continued collaboration of the DataWorkplace and the realization of concrete insights and instruments for participating organizations already hint at the benefits of transdisciplinary research for researchers and parties outside the university. The remainder of this chapter reflects more systematically on such benefits – as well as the challenges posed by working in this way. This section considers these benefits and challenges under three pertinent rubrics: measuring their impact on academic researchers, on government organizations, and on networked collaborators. For each of these categories we discuss our own broad learning points along with examples from our experience with the DataWorkplace, illustrating the benefits and challenges that arise within each category.

For Academic Researchers

For academic researchers, the move from traditional research methods to transdisciplinary research indeed generates both benefits and challenges. For each research project undertaken at the DataWorkplace, the process followed a similar pattern. First, student researchers as well as senior researchers had extended contact with DataWorkplace partner organizations to gain an understanding of the work processes and problems within a particular organization. This introductory phase was followed by close collaboration with mid-level practitioners to design the research, as well as a period of observation and data collection. The next stage involved the translation of these research results into practical recommendations or the development of an instrument aimed at solving or preventing the practical problem that had been identified. The final phase was devoted to bringing this contextual knowledge to a broader group of interested practitioners in and beyond the consortium.

For example, in developing the Meaningful Digital Collaboration Tool with the province of South Holland, the student researcher started the research process by shadowing ten of the organization's members to understand the

digital transformation taking place in the organization (Nieuwenhuizen 2020). This process was disrupted by the Covid-19 pandemic, necessitating that all provincial workers had to work from home. Together the researcher and practitioners discovered that working from home impoverished their work processes: working online made collaboration, contact, and creativity more difficult. Based on this finding, they developed the Meaningful Digital (and Hybrid) Collaboration tool, which teams could use to discuss the challenges of working digitally and thereby develop team-specific solutions.

This process reveals that one of the greatest benefits for researchers is the possibility of being immersed in the area being investigated; researchers work closely with practitioners and thus benefit from extremely rich and detailed data from within organizations. They also gain knowledge about the practices taking place in these organizations. The process of collaboration with practitioners also delivers additional insights and ways of approaching problems because of the knowledge contributed by practitioners to the development of the research, which is also a benefit. In the case of the dashboards, the research pointed the practitioners to the limited degree of literacy in understanding the displayed information: the intended purpose of the dashboard was never explicitly communicated. There were nonetheless expectations that the dashboard could solve policy problems and carry out administrative processes. Here, a need for demystification was directly brought forward by the practitioners. In addition, the collaboration provided very granular insights into the power asymmetries, different agendas, and discourses within the organization, and how these affect the response to new technologies and their implementation.

For the junior researchers, usually master's students from governance, media studies or applied ethics, working within the DataWorkplace created rewarding learning opportunities. Students were able to write their respective theses related to an actual case study and they acquired hands-on experience conducting research in the field and collaborating with practitioners (e.g., Nieuwenhuizen 2020; Ettlenger 2021). Alumni of the DataWorkplace frequently obtain PhD research positions. Others get jobs at consulting firms or in government organizations directly after graduating.

As rewarding as transdisciplinary work is for researchers in terms of data and experience, this sort of work also poses challenges for researchers. One such challenge involves the depth of the research. Student researchers and project-based research often operate in cycles of a few months, perhaps too short a span to achieve sufficient depth. Practitioners in these organizations sometimes mention that the turnover of researchers within

the DataWorkplace as well as the short timeframe can be obstacles to the achievement of long-term outcomes in their organizations.

Another challenge for academic researchers concerns bias. Because of the co-collaboration process, transdisciplinary research in the DataWorkplace fosters strong connections between researchers and practitioners in these organizations. In principle these connections are not negative, but close relationships between researchers and their contacts in the organization can make it trickier to remain objective while carrying out the research. While we did not observe that such closeness affected the quality of the research, we noticed that it occasionally led to an awkward blurring of roles. In one instance, a junior researcher reported finding herself helping a civil servant to organize an event – looking up addresses, checking schedules, and sending emails – though the event was not related to the research project.

Finally, the largest challenge for academic researchers in conducting transdisciplinary work is systemic: the academic system rewards publishing for one's academic peers, not for making societal impact. Transdisciplinary work, however, is strongly focused on impact, on delivering practical – and often local – solutions. Precisely because this type of work differs from traditional methods of scientific work, we find it more difficult to fit ourselves into the traditional evaluation of academic performance, which is concentrated on publishing research in high-impact journals. Although the international research community agrees that all data are situated in local contexts (Loukissas 2019), investigating contexts outside the English-speaking world is nonetheless a disadvantage. There is an additional burden of translating research which unfolded in a local language into English, and often to convince journals that it is even worth publishing.

The practice of the DataWorkplace may actually be leading the way towards new forms of evaluating academic performance that emphasize societal value more than academic metrics lacking meaning in the 'real world' (see also Utrecht University 2021; Miedema 2022).

For Government Organizations

Government organizations have shown real enthusiasm and motivation to engage in transdisciplinary work with researchers. Given the proliferation of data and the complexity of the societal problems these organizations hope to solve, these entities recognize the need to change their organizational processes. Engaging in transdisciplinary research can be a valuable resource for this sort of transition for several reasons.

For the DataWorkplace's partner organizations, the practical recommendations for addressing specific issues they are struggling with, and the instruments developed in the process, are among the most prominent benefits of this transdisciplinary work. Members of these organizations have been very enthusiastic about the practical knowledge developed in conjunction with the DataWorkplace, describing the DataWorkplace's uniqueness in its ability to deliver "fantastic" concrete products. Although these organizations are focused on the development of practical knowledge, prolonged collaboration between practitioners and researchers yields a host of other benefits.

First, we see an added benefit when the approaches taken by practitioners and researchers towards problems are combined. Practitioners, often saddled with busy workloads and many tasks to complete, frequently focus on finding solutions – a manifestation of a 'what works' mentality. In contrast, academic research adopts a different orientation, focusing on the system as a whole and taking more time to examine a given problem and its background. When perspectives are combined as in transdisciplinary collaboration, an organization's established practices can be looked at with fresh eyes. Through involvement with the DataWorkplace, informal practices and experiential knowledge were developed, informal networks of data-savvy staff emerged across different departments, and there were even formal training programs set up to teach and advance digital-ethical literacy. Practitioners frequently cited the benefit of taking the time to reflect on their practices with researchers and developing new ways of understanding socio-technical dynamics, which add real value for their organizations.

Despite these benefits for the organizations participating in the DataWorkplace, transdisciplinary research also presents significant challenges. From our experience, organizations have difficulty transforming what they have learned into lasting organizational change. This was particularly true with regard to implementing the instruments into organizational processes. Conversations and interviews with practitioners provided us with very positive feedback on the instruments. And yet very few of the instruments – despite being scientifically supported and viewed positively by the participating organizations – ended up being used on a long-term basis within the organizations.

This finding highlights how challenging it is to implement the outcomes of transdisciplinary work. We as researchers were prompted to reflect on the central role of instruments in our projects, which led us to decide to take a broader perspective on the relation between contextual and generic insights. This also highlighted how challenging it is for the participating organizations to use research outcomes and integrate them into their organizations.

Transdisciplinary work asks that organizations adopt a certain mentality shift – transdisciplinary research does not deliver readymade solutions but requires work on the part of the organization. Given that these government organizations are so caught up in their day-to-day routines, they must find ways to free up capacity (e.g., time, expertise, authority) not only to develop transdisciplinary measures but also to implement them. Some organizations may find it difficult to be open to the sort of critical reflection that can take place in collaborative research.

Finally, one of the foremost challenges of collaboration between academics and government organizations in a transdisciplinary way lies in developing research agendas that are both societally and academically relevant. We see that the participating organizations have very practical problems within their organization, and they hope that collaboration with the DataWorkplace can solve them. However, if the focus swings to solving such practical problems, it is often not broad or theoretical enough to be academically relevant. Finding this balance between societally and academically relevant research agendas requires constant attention and reflection.

For the Network of Government Organizations

One of the unique characteristics of the DataWorkplace is that it is organized as a network: it is not merely a transdisciplinary collaboration between the university and a single government organization. The DataWorkplace creates benefits because of the multiple options it provides for learning, knowledge dissemination, and cooperation across different organizations. But the full utilization of these options requires attention and effort.

The DataWorkplace provides multiple layers on which learning can take place. For government organizations, it occurs at the individual, organizational, and strategic levels. There is also a networked layer of learning. Here, organizations within the DataWorkplace learn from one another and connect to other participants and organizations from their respective networks. Via this expanded network, they learn about other organizations' experiences in dealing with datafication's various challenges as well as the research being conducted within the various organizations.

To give a more concrete example of the different levels of learning that transpires in the DataWorkplace, we can cite the development of the Data Team Start instrument. As a result of this tool's development at the municipality of Gouda, individual employees described learning about potential challenges associated with data projects. At the team level, they described learning about a new process for starting data projects. This learning also

extended to the strategic level of the organization, which incorporated the use of DataWorkplace instruments into its vision for data-driven work. Several other organizations would also use the insights developed through Data Team Start to re-think data competencies in their organizations, which shows the network learning that is taking place.

These opportunities for learning are all very positive, but it is nonetheless challenging for organizations in the DataWorkplace to realize these learning outcomes at the various levels. Organizations can find it challenging to be open to other organizations' experiences. Sharing experiences and knowledge with other organizations costs time and the benefits may not be directly tangible. Thus we note a tendency for knowledge to be shared between the university and one partner organization rather than being shared among the network. Although most of the partner organizations joined this collaboration to reap the benefits of learning from each other, we see that, in practice, such networked learning is more difficult to achieve.

Conclusions

Overall, the DataWorkplace presents a unique opportunity for collaborative investigation with practitioners into the impact of datafication on government organizations. Furthermore, our research not only aids these organizations in addressing challenges posed by technological change but also significantly transforms our own work as researchers. Instead of developing research merely for yet another peer-reviewed publication, we now have a different process of generating knowledge and disseminating it. This approach might lead to fewer traditional academic publications, but it facilitates more effective knowledge transfer to the sectors involved.

Our experience with the DataWorkplace highlights the benefits of transdisciplinary research. For academics, transdisciplinary research provides access to and firsthand learning about local and regional government organizations. For government organizations, transdisciplinary research introduces a new way of reflecting on their own practices and enables the application of research findings to improve their practices and further organizational change. For networked collaborators, it provides an environment for community-engaged learning at multiple levels: within academia, within each organization, and within the network of participants. Finally, the development of new knowledge and practical instruments for government organizations to respond to the complex challenge of datafication is something that benefits society at large.

Transdisciplinary collaborative research is not without its challenges. The environments of academia and government organizations can be difficult contexts for such efforts to flourish. At the same time, in the DataWorkplace we developed strategies to tackle these challenges and were able to create value for academia and for local and regional government organizations. Working on achieving and sustaining three types of balance – for researchers: between realizing local impact but also contributing to generic academic knowledge; for government organizations: between focusing on key tasks and being open to transdisciplinary learning and organizational change; for the joint collaboration: between research agendas that are both societally and academically relevant – demands constant attention and reflection.

Overall, this case highlights that transdisciplinary research requires curiosity, perseverance, empathy, and the courage to leave established trails behind and to travel new roads of joint knowledge creation.

References

- Branderhorst, Martiene. 2020. "Digitaal Leiderschap: Verkenning van de Veranderende Rol van Gemeentesecretarissen in de Informatiesamenleving." *Bestuurswetenschappen* 74 (4): 26–43. <https://doi.org/10.5553/Bw/016571942020074004003>.
- Das, Djurre, Patricia Faasse, Bart Karstens, and Paul Diederer. 2020. *Raad Weten met Digitalisering: Hoe de Gemeenteraad Kan Sturen op de Maatschappelijke Impact van Digitale Technologie*. The Hague: Rathenau Institute.
- Ettlinger, Krista M. 2021. *Transdisciplinary Social Learning and Transdisciplinary Capacity: Responding to the Challenge of Transdisciplinary Knowledge Integration in the Public Sector*. Master's thesis, Utrecht University School of Governance.
- Franzke, Aline, Iris Muis, and Mirko Tobias Schäfer. 2021. "Data Ethics Decision Aid (DEDA): a dialogical framework for ethical inquiry of AI and data projects in the Netherlands." *Ethics and Information Technology* (2021): 1–17.
- Grimmelikhuijsen, Stephan, and Albert Meijer. 2022. "Legitimacy of Algorithmic Decision-Making: Six Threats and the Need for a Calibrated Institutional Response." *Perspectives on Public Management and Governance* 5 (3): 232–242.
- Jahn, Thomas, Matthias Bergmann, and Florian Keil. 2012. "Transdisciplinarity: Between Mainstreaming and Marginalization." *Ecological Economics* 79 (July): 1–10. <https://doi.org/10.1016/j.ecolecon.2012.04.017>.
- Jansen, Julia, Sue-Yen Tjong Tjin Tai, Jos van den Broek, and Jasper Deuten. 2020. "Stad Zoekt Toga – Universiteiten en Hogescholen als Structurele Kennispartner voor Gemeenten." The Hague: Rathenau Institute.

- Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. 2014. "Big-Data Applications in the Government Sector." *Communications of the ACM* 57 (3): 78–85. <https://doi.org/10.1145/2500873>.
- Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Los Angeles, CA: Sage.
- Klein, Julie. 2004. "Prospects for Transdisciplinarity." *Futures* 36 (4): 515–26. <https://doi.org/10.1016/j.futures.2003.10.007>.
- Klievink, Bram, Bart-Jan Romijn, Scott Cunningham, and Hans de Bruijn. 2017. "Big Data in the Public Sector: Uncertainties and Readiness." *Information Systems Frontiers* 19 (2): 267–83. <https://doi.org/10.1007/s10796-016-9686-2>.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT press.
- Maciejewski, Mariusz. 2017. "To Do More, Better, Faster and More Cheaply: Using Big Data in Public Administration." *International Review of Administrative Sciences* 83 (1_suppl): 120–35. <https://doi.org/10.1177/0020852316640058>.
- Meijer, Albert. 2018. "Datapolis: A Public Governance Perspective on 'Smart Cities'." *Perspectives on Public Management and Governance* 1 (3): 195–206.
- Meijer, Albert, Mirko Tobias Schäfer, and Martiene Branderhorst. 2019. "Principes Voor Goed Lokaal Bestuur in de Digitale Samenleving: Een Aanzet tot een Normatief Kader." *Bestuurswetenschappen* 73 (4): 8–23.
- Meijer, Albert, and Erna Ruijter. 2021. *Code Goed Digitaal Openbaar Bestuur*. Government of the Netherlands. <https://www.rijksoverheid.nl/documenten/rapporten/2021/04/30/code-goed-digitaal-openbaar-bestuur>.
- Meijer, Albert, Alex Ingrams, and Stavros Zouridis. 2022. *Public Management in an Information Age: Towards Strategic Public Information Management*. London: Bloomsbury Publishing.
- Micheli, Marina, Marisa Ponti, Max Craglia, and Anna Berti Suman. 2020. "Emerging Models of Data Governance in the Age of Datafication." *Big Data & Society* 7 (2): 205395172094808. <https://doi.org/10.1177/2053951720948087>.
- Miedema, Frank. 2022. *Open Science: the Very Idea*. Dordrecht: Springer.
- Ministerie van Infrastructuur en Milieu. 2016. *Omgevingswet in het Kort: Ruimte voor Ontwikkeling, Waarborgen voor Kwaliteit*. Interdepartementale Programmadirectie Eenvoudig Beter. www.omgevingswetportaal.nl.
- Mobjörk, Malin. 2010. "Consulting versus Participatory Transdisciplinarity: A Refined Classification of Transdisciplinary Research." *Futures* 42 (8): 866–73. <https://doi.org/10.1016/j.futures.2010.03.003>.
- Nieuwenhuizen, Wouter. 2020. *How Working Digitally During the Covid-19 Pandemic Does Not Help to Transform the Public Sector: A Digital Organisational Ethnography*. Master's thesis, Utrecht University.

- Redden, Joanna. 2018. "Democratic Governance in an Age of Datafication: Lessons from Mapping Government Discourses and Practices." *Big Data & Society* 5 (2): 205395171880914. <https://doi.org/10.1177/2053951718809145>.
- Ruijter, Erna. 2021. "Designing and Implementing Data Collaboratives: A Governance Perspective." *Government Information Quarterly* 38 (4): 101612.
- Ruijter, Erna, Justien Dingelstad, and Albert Meijer. 2023. "Studying Complex Systems Through Design Interventions Probing Open Government Data Ecosystems in the Netherlands." *Public Management Review* 25 (1): 129–149.
- Schmidt, Laura, Thomas Falk, Marianna Siegmund-Schultze, and Joachim H. Spangenberg. 2020. "The Objectives of Stakeholder Involvement in Transdisciplinary Research. A Conceptual Framework for a Reflective and Reflexive Practise." *Ecological Economics* 176 (October): 106751. <https://doi.org/10.1016/j.ecolecon.2020.106751>.
- Schäfer, Mirko Tobias, Karin van Es, and Iris Muis. 2023. "Investigating the Datafied Society. Entrepreneurial Research as Approach." In *Situating Data: Inquiries in Algorithmic Culture*, edited by Karin van Es and Nanna Verhoeff, 267–272. Amsterdam: Amsterdam University Press.
- Siffels, Lotje, David van den Berg, Mirko Tobias Schäfer, and Iris Muis. 2022. "Public Values and Technological Change: Mapping How Municipalities Grapple with Data Ethics." In *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, edited by Andreas Hepp, Juliane Jarke, and Leif Kramp, 243–265. Cham: Palgrave Macmillan.
- Utrecht University. 2021. "Recognition and Rewards Vision." University policy. Utrecht: Utrecht University. <https://www.uu.nl/sites/default/files/UU-Recognition-and-Rewards-Vision.pdf>.
- Van Es, Karin, and Mirko Tobias Schäfer. 2017. "New Brave World." In *The Datafied Society. Studying Culture through Data*, edited by Mirko Tobias Schäfer, and Karin van Es, 13–22. Amsterdam: Amsterdam University Press.

About the Authors

Krista Ettliger is a PhD Candidate at Utrecht University School of Governance in the Netherlands. Her research interests include collaborative innovation, democratic innovations and forms of citizen participation. Her PhD research studies the combination of deliberative and plebiscitary forms of democratic innovations and their potential contribution to the revitalization of representative democracy.

Mirko Tobias Schäfer is Associate Professor of AI, Data & Society at Utrecht University's research area *Governing the Digital Society* and the Department

for Information & Computing Sciences. Mirko is co-founder and Sciences lead of the Data School. He studies the datafication of public management and engages in the development of responsible and accountable AI and data practices.

Albert Meijer is a professor of public innovation at Utrecht University in the Netherlands and chair of the Public Management and Governance Group. He is a well-known expert in public administration in an information age with publications on public innovation, smart cities, ICT, data and social media in the public sector.

As managing director, **Martiene Branderhorst** is involved in making the outdoor space of The Hague better in a sustainable way. Martiene has written a dissertation on the use of management information and is currently conducting research at Utrecht University on the consequences of digital technologies for the role of public managers.

15. You Will Be Assimilated: Reflections on Ethnographic Fieldwork on Algorithmic Systems

Daan Kolkman

Abstract

For over a decade, I have navigated the dual identities of being both a researcher and a practitioner within the data science domain. This unique vantage point has provided me with invaluable insights into the daily workings and challenges faced by those who develop and deploy algorithmic systems. However, this dual role has not been without its tensions. In this chapter I explore these tensions and consider if – and if so, how – critical algorithm studies can be pursued while being embedded as a data professional. More specifically, I reflect on my own gradual ‘assimilation’ in the community of data professionals or my transition from playing an outsider to an insider role in my fieldwork on algorithmic systems.

Keywords: Ethnography; Expertise; Artificial intelligence; Sociology

Introduction

Prior to the early 2000s, the term ‘data scientist’ would have sounded outlandish to all but a very small contingent of statisticians and computer scientists. Although Peter Naur famously coined the term in 1974, it did not become common parlance until three decades later. The chess victory of IBM’s Deep Blue over Gary Kasparov in 1997 must have sparked the imagination of some, but machine learning and artificial intelligence were fringe fields before the turn of the millennium. As recent breakthroughs pertaining

to Large Language Models (LLMs) show, to say that data-centric disciplines have evolved over the last few decades is something of an understatement.

The algorithmic systems that embody this evolution have been heralded by some as tools that may help solve some of our most pressing societal problems (Floridi and Taddeo 2016) due to the superior performance and objectivity of these tools (Kolkman 2020). Yet an increasing number of incidents show that algorithmic systems are fallible and can be biased (Angwin et al. 2016); they can also facilitate exploitation (van Doorn 2017) and perpetuate inequality (Eubanks 2018). These incidents have sparked calls for further research and regulation. Nevertheless, the fallibility of algorithmic systems is hardly a new phenomenon, with authors as early as Black (1962, 225) pointing out that the “attendant dangers are [...] obvious.” Moreover, prominent mishaps with algorithmic systems have occurred since we started using computers (see Pielke 1999; Van der Sluijs 2002).

It seems, then – even with the EU AI Act on the horizon – that we have not learned from our mistakes, despite efforts to develop guidelines for the proper implementation of algorithms in high-stakes decision-making contexts (Dekker, Groenendijk, and Sliggers 1990; Treasury 2015). As a sociologist studying algorithmic systems, I recognize the urgency of advocating for more critical empirical research into the communities of data professionals (Gillespie 2014). However, it is equally important that we introspect and learn from our own efforts studying algorithms in practice (among others Christin 2017; Kolkman 2020; Seaver 2017; Young et al. 2019). Failure to do so would not only hinder our understanding of algorithmic systems but might also cause us to fall into the trap of unlearned lessons, something we critique when others fall short in this regard.

Fortunately, we find such introspection from Seaver (2017), who shares his ‘tactics’ for the study of algorithmic systems, and Jatón (2021), who offers an in-depth discussion of the foundations of his ethnographic methods. My approach, while similar, took a slightly different point of departure. This may explain in part why I did not observe the practices of secrecy, recklessness, or ill intent that others have described (O’Neil 2016; Pasquale 2015). Specifically, I built on studies by Beunza and Garud (2007), Beunza and Stark (2012), and Spears (2014) that – like those of Seaver and Jatón – center on understanding the socio-material context of algorithms. However, such work in the Social Studies of Finance domain also focuses on understanding the ‘technical’ aspects of algorithms themselves.

In this chapter, I share key lessons from a decade of studying algorithms in practice. My journey, which started in 2012 with a PhD focused on computational models in government, offers a unique lens to examine the

intricacies of accessing algorithmic systems and of transitioning from an observer to a participant in the field, as well as to engage with the insights and ethical considerations that emerge from such a transformation.

Embarking on an Ethnography of Algorithms

From the outset of my PhD, I was driven to understand more about why models were considered useful tools in government (see Kolkman 2016). My interest centered specifically on when and why the use of models beneficially – or adversely – impacts decision-making. Given this aim, a qualitative research design seemed a natural choice. A variety of qualitative research designs and methods were suitable for the study of the digital as it unfolds in practice (see Roberts et al. 2016). Intrigued by the laboratory studies conducted by Science and Technology Studies scholars in the 1970s (Sismondo 2010), I opted for an ethnographic approach. Through the idea of “anthropological strangeness” (Latour and Woolgar 1979, 29), according to which a lack of prior knowledge does not prevent understanding of the object of study, these laboratory studies facilitated a deeper understanding of what exactly scientists do in laboratories (Shapin and Schaffer 1985).

I sought to develop rich accounts with a focus on descriptive detail (Hammersly and Atkinson 1995) and intended to explore how power is manifested through practices, privilege, and the normalization of specific types of knowledge (Jain and Jadhav 2009). Following Marcus’ (1995) suggestions on studying communities of professionals, I opted for a multi-sited ethnography consisting of eight case studies in organizations where ‘computational models’ were used to inform decision-making.

Inspired by Social Studies of Finance, I did not adopt the ‘anthropological strangeness’ perspective as radically as had scholars in the laboratory studies, but I certainly considered myself, with regard to the communities of data professionals I was studying, to be an outsider (Adler and Adler 1987). In Gold’s (1958) classic typology of research roles that range from complete observer to complete participant, I was somewhere between complete observer and observer as participant. I had some postgraduate exposure to statistics, geographic information systems, and modelling; however, I completely lacked domain knowledge about – or statistical expertise on – the types of computational models that were used in government. Indeed, one finding of my work is that data professionals working on a particular model – or algorithm – may inhabit a distinct intellectual community with specific modelling practices and associated tacit knowledge (Kolkman 2022). Approaching my case studies

and the data professionals as an outsider came with a set of interwoven challenges. I discuss some of the more prominent challenges below.

Getting Access and Studying Up

It can be problematic for an outsider to any community to negotiate access for research (Dwyer and Buckle 2009). While Seaver (2017) points out that resistance to accessing algorithmic systems can be thought of as data in its own right, I wanted to get closer to my research subject. For data professionals, participation in a qualitative study on work practices surrounding algorithmic systems comes with risks. This was particularly true in the UK at that time (the Laidlaw Inquiry (2012) had brought modelling into the public eye) and is perhaps even more true today, since we have now been exposed to many episodes of algorithmic bias. From the perspective of a data professional, it requires a considerable degree of trust to meet with an outsider for an interview, to allow that outsider to sit in on meetings, and to permit that outsider access to their algorithms. There had to be trust that my stated intentions were true and that I was not seeking to expose another model mishap or case of algorithmic bias. I vividly remember an instance when a data professional seemed particularly distrustful:

So, can you tell me a bit more about where you are coming from? Are you some sort of research journalist that's seeking to uncover dirt? We have had people like that before and all they did was see conflict where there really wasn't any to begin with.

By allowing access to their organizations, the data professionals were making an implicit endorsement of me as an outsider and thus were putting their reputations at risk. In negotiating access to organizations where algorithms were developed, I benefited from the sponsorship of my PhD committee. Nonetheless, on two occasions I was asked to sign a confidentiality agreement stipulating that I could not publish any of my findings for five years. These confidentiality agreements were signed by the university and by me. As I was at that time an early career researcher facing the pressure to publish or perish, these stipulations were – to put the matter most charitably – an inconvenience. However, access to these communities was hard to come by and so, following the guidance of my supervisors, I was keen to oblige (see Petre 2021 for a similar experience).

To the data professionals, my pursuit of the study as part of my PhD may have demonstrated my seriousness and might have ensured that I

could be trusted with access to sensitive information. The confidentiality agreements helped cement this perception in the cases where additional safeguards were considered necessary. Being a PhD student, I often found myself in a situation known in anthropology as “studying up” (Nader 1972), with the researcher having lower status and less power than the subjects under study. In my case, this power imbalance produced various small indignities: I was frequently kept waiting for appointments, had to deal with data professionals not showing up for interviews, was mistaken for a delivery person, and had to cope with participants clearly doing something else on their phone during interviews.

Reflecting on my interactions, I am fairly certain I posed no threat to the model professionals I worked with. If anything, my presence as an academic may have given certain model professionals a status boost: I sat in on meetings where I was introduced as an expert from academia involved in reviewing the computer model that was the subject of the meeting. I thus sometimes felt enlisted in the performance of some sort of symbolic capital by the data professionals. This dynamic is akin to what Abidin (2020) describes in her study of the influencer industry. Although her research context – focusing on social media influencers – appears distinct from government algorithmic systems, she, too, experienced being perceived as a ‘trophy acquaintance’, a role that added value to the social standing of her subjects.

Establishing Rapport and Developing Expertise

To establish rapport and build trust with the data professionals, and to ensure they took me seriously, I felt I had to project confidence and demonstrate expertise with regard to my participants’ data-intensive work. This feeling was confirmed on numerous occasions during the exploratory phase of my research, for example when some of the data professionals asked about my educational background or tried to engage me in a discussion about some statistical or modelling concept such as sensitivity analysis, uncertainty bounds, or calibration. I soon found that most model professionals would dismiss or ignore my further invitations if I could not connect with them on the grounds of their subject-matter expertise. The problem with developing such expertise, however, is finding a place to start (Spears 2014).

One strategy is to try and develop what Collins (2004) refers to as “inter-activeness”. This is a type of expertise through which one can interact meaningfully with members of a particular community, while not possessing the tacit knowledge required to be a contributing member of that

community (Collins and Evans 2007). In the case of model professionals and their algorithms, this means developing knowledge of, among other things, the specific models they use, their logic and structure, and the various problems they face in implementing these models within their organizations. What such interactional expertise does not entail is mastering the largely tacit knowledge needed to contribute to the community itself: the skills needed to design, develop, and use algorithms.

By definition, interactional expertise flows from interaction with insiders and so – perhaps somewhat ironically – the most effective way to acquire this type of expertise is through immersion in the community one seeks to study. This very immersion is what I was looking for in the first place. In the context of studying the practices of data professionals working on models in government, this would have entailed substantial time spent working with model professionals in different organizations. In that regard, I was constrained by the three-year duration of my research funding. The difficulty of developing this expertise was further exacerbated by the variety of domains represented by the models I was studying. So I instead pursued a substantial amount of self-study, reading, on my own time, textbooks and research papers on a variety of topics such as environmental modelling and pensions modelling. Through this effort I became familiar with what Collins and Evans (2007, 14) call “primary source knowledge”. I continued this process of ‘off-site’ learning throughout the research project when new concepts were brought up by participants. This sort of commitment, along with my many discussions with data professionals, helped build my credibility, but I nevertheless did not learn the tacit knowledge unique to a community.

Tacit Knowledge and Data Science Tools

How, then, does one acquire tacit knowledge before entering the field? The short answer is that it is virtually impossible to do so, given that data professionals form *epistemic communities* (Knorr-Cetina 1981). The practices, privilege, and the normalization of specific types of knowledge cannot be learned off-site, and neither can the expertise that stems from operating a specific model. Data professionals make use of many digital tools in their day-to-day work, and these tools may require tacit knowledge to operate. Familiarity with these tools is essential not only to interact with data professionals but also to explore the digital dimension of their work and communities (Góralaska 2020). Given that many of these tools are used not just in one organizational setting but across communities of data professionals, it was possible for me to develop tacit knowledge while off-site.

Some of the tools will be familiar to anyone who has been an undergraduate: Microsoft Office, SharePoint, or the Slack messaging system. Other tools more specific to the data profession included statistical software such as SAS or SPSS, programming languages such as R, Python, or Fortran, and versioning tools such as Git. Since it was impossible for me to familiarize myself with all these tools, I focused on those in the Python ecosystem. Nonetheless, I found it was important to get some basic knowledge of the other tools to better understand the work that model professionals engage in. Specifically, I brushed up on my statistics knowledge and developed some basic software engineering know-how through several postgraduate, Coursera, and Udacity courses. To acquire as much of the tacit knowledge that a data scientist would have acquired, I approached these courses as if I were studying to become a data scientist or statistician.

Assimilating as a Data Scientist

At the outset of my PhD, I felt like an outsider with regard to the communities of data professionals I was studying, but my role started to change towards the end of the project. Over the course of two-and-a-half years of fieldwork and study, I had developed considerable expertise in modelling and data science methods. After my funding ran out, I started taking on some freelance data analysis work. At first, this work built mostly on the statistical expertise I had acquired for my undergraduate and postgraduate degrees. However, I soon found that the knowledge of subjects such as machine learning – which I encountered in my fieldwork – was of particular interest to many organizations. I decided to enroll in an intensive summer course aptly titled “Science to Data Science”. Gradually, I moved beyond the interactional expertise level and developed a decent understanding of a modest selection of data science and machine learning methods.

My newfound status as a ‘data scientist’ afforded me a steady income, but it also fundamentally impacted my role as a researcher. When I spoke to data professionals for my research, I was no longer ‘studying up’. I felt like I was more their equal and had transitioned to ‘studying sideways’ (Nader 1972). This transition came with considerable advantages. By virtue of my data expertise, I was taken more seriously, but I also got “primary access” (Bran-nick and Coghlan 2007, 67) to several algorithmic systems. More generally, it became easier for me to access settings in which algorithms were developed and to establish rapport and trust with data professionals. However, being an insider also brought several new and distinct disadvantages. As Labaree

(2002, 12) points out, any researcher in an insider position must beware of “ethical and methodological dilemmas”. Below, I describe such challenges, which I found – and still find – difficult to navigate.

Complicity

Since I could no longer consider myself fully outside the community of data professionals, I felt less and less at ease with stark criticisms of their work, but at the same time I did not feel comfortable defending their work. Earlier on, I found that this group of professionals worked hard and diligently to supply information to senior decision-makers in government. Nonetheless, I saw that they sometimes struggled to keep up with the latest quality-assurance guidelines and decision-making time pressures. What I observed in my studies, however, were not the practices of secrecy, recklessness, or ill intent that others have described (O’Neil 2016; Pasquale 2015). I am not questioning the importance of such work: it seems to me the deeply problematic nature of some algorithms has been demonstrated beyond doubt. Yet I feel the discussion would have benefitted from some nuance and a different perspective on the people who build algorithms. However, I felt that I had little moral authority to voice such a view because of my assimilation as a data professional.

Ghorasi and Wels (2009) suggest that no member of society at large can escape complicity in the perpetuation of a variety of injustices, and that ethnographers are no exception. Of course, this does not absolve ethnographers from responsibility. Rather, they have a pivotal role to play in directing attention away from the “algorithmic drama” (Ziewit 2016) perpetuated by proponents and opponents of machine learning algorithms.

Ethnographers can reject the arrogance of ‘the moral high ground’ (Ghorasi and Well 2009, 244) and adopt an active mediating role or engage in reconciliatory discourse. They are well positioned to do so because of their specific temporality, which places them ‘betwixt and between’ (Turner 1969). In my case studies, I have been embedded within a group of data professionals for at most a year, and the same goes for my work as a freelance data scientist. Yes, this liminality comes with a loss of impartiality, but it might also position me as being just acceptable enough to scholars of algorithms and data professionals alike.

From my own experience, such a role comes with a certain feeling of being disconnected or even anxious (see Blomley 1994). While the insider perspective can be very valuable in developing a deeper understanding of algorithmic systems, I am perhaps no longer distanced enough from data

professionals to function as a moral counterweight. Thus I feel vulnerable to suggestions that the standard of my research could be questioned or its validity threatened (see Stanley and Wise 1993). At the same time, my preoccupation with qualitative research does not permit me full membership in the data science profession either, as these professionals may rightly view my research as deeply qualitative and at odds with their data-driven worldview.

Covert Research?

Another issue arose with my becoming an insider: the reason I had access to the data professional communities had little to do with research. These organizations are interested in working with me because of my knowledge of data science, not necessarily because of my research on data science practices. Although I talk about my qualitative work and let those I work with know when I intend to publish something in a journal, this is of little concern to them. In practice, this situation entails a shift from a more overt to a more covert research approach. Anderson and Bissel (2004) point out that in practice, there is not always a clear-cut distinction between the two extremes of this continuum, an observation confirmed by my experience. They provide the comforting suggestion that the practice of building rapport could also be interpreted as “encourag[ing] people to forget that you are constantly observing them and registering everything they are saying and doing” (Bourgois 2007, 296–97).

Regardless, covert qualitative research is an emotive approach and is thus controversial. The most common issue with this approach is the justification of deliberate deception (Calvey 2013). Bulmer (1982, 252), who has written extensively on the topic, usefully defines covert research as referring to research situations where the real identity of the observer as a social researcher remains entirely unknown to those with whom they are in contact. The investigator purports to be a complete participant and is in fact secretly something else.

Although I have not intentionally chosen a covert approach, it is hard to avoid it entirely. I probably talk more about issues of algorithmic bias, fairness, and transparency than others, but not everyone I talk to will be aware of my qualitative research efforts. Within what Calvey (2008, 3) calls the ‘standard view’ on covert research, there is a recognition that informed consent is not always achievable in its absolute form. With regard to such contexts there has been a call for the retrospective explicit debriefing of subjects, so that at least they know that research had taken place even

if they did not know this beforehand (Calvey 2008). Such debriefing can also be done in relation to a participant validation (Pighini et al. 2014) or triangulation schemes (Carter et al. 2014), in which findings are presented to the participants for feedback. This is the approach I have ventured to use in my current studies.

Career Progress

My oscillation between the data professional community and academia means that I am sometimes confronted with institutional boundaries and arrangements from both communities. While my work on the SME Datalab¹ – an initiative which connected more than 300 Master's students and SMEs – was welcomed by practitioners and policymakers, it has done little to advance my academic career. This is not an issue per se, but it did – and does – affect my career trajectory in academia. For instance, although a broad alliance of academic institutions launched a new 'Recognition and Rewards' initiative which puts less emphasis on the number of publications in high-impact journals and more on education and societal impact made by academics, the effects of this change in policy still need to percolate through university systems.

In my case, I am engaged with many of the activities of a typical academic: I am involved in teaching, I do research, I join departmental working groups and discussions. Since finishing my PhD, however, it was not until very recently that I held a formal academic position. After years of being employed as 'support staff' I joined the computer science department at Utrecht University in the role of thesis coordinator for their Applied Data Science Master's program. My expertise as someone who moves freely within and outside the university has prepared me well to build these bridges and foster the mutual knowledge transfer among the university, its researchers and students, and different organizations in the Netherlands and abroad.

Concluding Remarks

In this chapter, I have shared my experiences with conducting ethnographic fieldwork on data professionals and the algorithms they help develop. A primary takeaway from my experience is the indispensable value of an insider's perspective in conducting meaningful and impactful research in the

1 See <https://www.jadsmkbbdatalab.nl/>

field of algorithmic systems. This perspective adds a layer to our understanding, but it may also hold the potential to transform it by integrating the lived realities and tacit knowledge of practitioners into academic discourse.

To those who seek to undertake similar work, I would recommend the following. First and foremost, build meaningful relationships within the community of practice. This involves not just securing access but nurturing trust, demonstrating respect for the community's knowledge, and actively collaborating with them to address shared concerns. As highlighted when discussing the challenges of gaining access, the importance of fostering genuine relationships was underscored in my interactions with data professionals, where establishing trust was paramount. In addition, you should seek to develop dual expertise by investing in both theoretical understanding and practical skills pertinent to the field. In the context of studying algorithmic systems, this may entail learning how to code or how to query databases. This dual expertise allows you to engage deeply with their subject matter, fostering richer insights and more nuanced interpretations.

My secondary takeaway is that immersion in the sector over time leads inevitably to assimilation – which presents a challenge. While the insider perspective is invaluable, it also means you are no longer a distant observer. You need to try to maintain academic rigor. The complexities I encountered when my role as a researcher began to merge with that of a data professional illustrate the continuous need to reflect on one's positionality, acknowledging potential biases and rigorously applying methodological frameworks when analyzing data and interpreting findings.

If universities are serious about this type of research, they should reward researchers who seek to act as a conduit between academia and practice. This involves not just translating academic findings into practical interventions but also bringing real-world challenges and insights into academic discussions. This dual role enriches both realms, as seen in my own experience in starting the SME Datalab and transitioning from a PhD student to a data professional. Moreover, universities need to ensure the presence of robust support systems, including administrative, legal, and ethical frameworks, in order to facilitate this kind of engaged and immersive research. As immersed research, this involves working closely with support staff and requires advocating for institutional policies and funding structures that recognize and reward this engaged approach to research.

Undoubtedly, this hinges significantly on institutional backing. Given the often static nature of institutions, fostering an environment conducive to this sort of engaged research necessitates a concerted effort. It is through this collaborative endeavor that we can foster a fertile research landscape,

one where multifaceted and deeply engaged inquiry not only flourishes but also substantially enriches our comprehension and shaping of algorithmic systems. As the importance of these systems grows, this effort is important if we want to build algorithmic systems that contribute to solving – rather than exacerbating – societal problems.

References

- Abolafia, Mitchel Y. 1999. "Markets as Cultures: An Ethnographic Approach." *Sociological Review* 46 (1_suppl): 69–85.
- Adler, Patricia. A., and Peter Adler. 1987. *Membership Roles in Field Research*. Vol. 6. Newbury Park: Sage.
- Anderson, Claire, and Paul Bissell. 2004. "Using Semi Covert Research to Evaluate an Emergency Hormonal Contraception Service." *Pharmacy World and Science* 26 (2): 102–106.
- Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." ProPublica, March 23, 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Beunza, Daniel, and Raghu Garud. 2007. "Calculators, Lemmings or Frame-makers? The Intermediary Role of Securities Analysts." *Sociological Review* 55 (2_suppl): 13–39.
- Beunza, Daniel, and David Stark. 2012. "From Dissonance to Resonance: Cognitive Interdependence in Quantitative Finance." *Economy and Society* 41 (3): 383–417.
- Black, Max. 1962. *Models and Metaphors: Studies in Language and Philosophy*. Ithaca, NY: Cornell University Press.
- Bourgois, Philippe. 2007. "Confronting the Ethics of Ethnography: Lessons From Fieldwork in Central America." In *Ethnographic Fieldwork: An Anthropological Reader*, edited by A. C. G. M. Robben & J. A. Sluka, 288–297. Hoboken, New Jersey: Wiley-Blackwell.
- Brannick, Teresa, and David Coghlan. 2007. "In Defense of Being 'Native': The Case for Insider Academic Research." *Organizational research methods* 10 (1): 59–74.
- Calvey, David. 2008. "The Art and Politics of Covert Research: Doing 'Situated Ethics' in the Field." *Sociology* 42 (5): 905–918.
- Christin, Angèle. 2020. "The Ethnographer and the Algorithm: Beyond the Black Box." *Theory and Society* 49 (5): 897–918.
- Collins, Harry. 2004. "Interactional Expertise as a Third Kind of Knowledge." *Phenomenology and the Cognitive Sciences* 3 (2): 125–143.
- Collins, Harry, and Robert Evans. 2008. *Rethinking Expertise*. Chicago, IL: University of Chicago Press.

- Crawford, Vincent. P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica: Journal of the Econometric Society* 1431–1451.
- Creswell, John W., Cheryl N. Poth. 2017. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: SAGE.
- Davenport, Thomas H., and D. J. Patil. 2012. "Data scientist: The Sexiest Job of the 21st Century." *Harvard Business Review* 90 (5): 70–76.
- Davenport, Thomas H. 2020. "Beyond Unicorns: Educating, Classifying, and Certifying Business Data Scientists." *Harvard Data Science Review* 2 (2). <https://doi.org/10.1162/99608f92.55546b4a>.
- Department for Transport 2020. "Report of the Laidlaw Inquiry." Available online at: <https://www.gov.uk/government/publications/report-of-the-laidlaw-inquiry>. Last accessed on January 28, 2024.
- Dekker, C. M., Groenendijk, A., Sliggers, C. J., and Verboom, G. K. 1990. "Quality Criteria for Models to Calculate Air Pollution." *Lucht (Air)* 90. Leidschendam: Ministry of Housing, Physical Planning and Environment.
- Dwyer, Sonya Corbin, and Jennifer L. Buckle. 2009. "The Space Between: On Being an Insider-Outsider in Qualitative Research." *International Journal of Qualitative Methods* 8 (1): 54–63.
- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St Martin's Press.
- Floridi, Luciano, and Mariarosaria Taddeo. 2016. "What is Data Ethics?" *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083): 20160360.
- Freeman, Christopher, and Luc Soete. 2009. "Developing Science, Technology and Innovation Indicators: What We Can Learn from the Past." *Research Policy* 38 (4): 583–589.
- Ghorashi, Halleh, and Harry Wels. 2009. "Beyond Complicity: A Plea for Engaged Ethnography." In *Organizational Ethnography: Studying the Complexities of Everyday Life*, edited by Sierk Ybema, Harry Wels, Dvora Yanow. 231–253. Thousand Oaks: Sage.
- Gold, Raymond. L. 1957. "Roles in Sociological Field Observations." *Social Forces* 36 (3): 217–223.
- Góralaska, Magdalena. 2020. "Anthropology from Home: Advice on Digital Ethnography for the Pandemic Times." *Anthropology in Action* 27 (1): 46–52.
- Gordon, Sue. 2004. "Understanding Students' Experiences of Statistics in a Service Course." *Statistics Education Research Journal* 3 (1): 40–59.
- Hammersley, Martyn, and Paul Atkinson. 1995. *Ethnography: Principles in Practice*, second edition. New York, NY: Routledge.
- Jain, Sumeet, and Sushrut Jadhav. 2009. "Pills That Swallow Policy: Clinical Ethnography of a Community Mental Health Program in Northern India." *Transcultural Psychiatry* 46 (1): 60–85.

- Jaton, Florian. 2021. *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. Cambridge, MA: MIT Press.
- Kolkman, Daan. A. 2016. "Models in Policy Making." PhD Thesis, University of Surrey.
- Kolkman, Daan. A., Paolo Campo, Tina Balke-Visser, and Nigel Gilbert. 2016. "How to Build Models for Government: Criteria Driving Model Acceptance in Policymaking." *Policy Sciences* 49 (4): 489–504.
- Kolkman, Daan. 2020. "The Usefulness of Algorithmic Models in Policy Making." *Government Information Quarterly* 37 (3): 101488.
- Kolkman, Daan. 2022. "The (In) Credibility of Algorithmic Models to Non-Experts." *Information, Communication & Society* 25 (1): 93–109.
- Knorr-Cetina, Karin D. 1981. "The Micro-Sociological Challenge of Macro-Sociology: Towards a Reconstruction of Social Theory and Methodology." In *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro-Sociologies*, edited by Aaron Victor Cicourel, and Karin D. Knorr-Cetina, 1–47. Boston: Routledge & Kegan Paul.
- Kuhn, Thomas. S. 1963. "The Function of Dogma in Scientific Research." In *Scientific Change: Historical Studies in the Intellectual, Social, and Technical Conditions for Scientific Discovery and Technical Invention, From Antiquity to the Present*, edited by Alistair Cameron Crombie, 347–369. New York, NY: Basic Books.
- Labaree, Robert V. 2002. "The Risk of 'Going Observationalist': Negotiating the Hidden Dilemmas of Being an Insider Participant Observer." *Qualitative Research* 2 (1): 97–122.
- Laidlaw, Sam. 2012. *Report of the Laidlaw Inquiry: Inquiry into the Lessons Learned for the Department for Transport from the Intercity West Coast Competition*. London: The Stationery Office.
- Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life*. Princeton, NJ: Princeton University Press.
- Nader, Laura. 1972. "Up the Anthropologist: Perspectives Gained from Studying Up." In *Reinventing Anthropology*, edited by Dell Hymes, 284–311. New York, NY: Vintage.
- Newquist, Harvey. P. 1994. *The Brain Makers*. Indianapolis, IN: Sams Publishing.
- Murthy, Dhiraj. 2008. "Digital Ethnography: An Examination of the Use of New Technologies for Social Research." *Sociology* 42 (5): 837–855.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Broadway Books.
- Patil, Dhanurjay J. 2011. *Building Data Science Teams*. O'Reilly Media, Inc.
- Pielke, Roger. A. 1999. "Who Decides? Forecasts and Responsibilities in the 1997 Red River Flood." *Applied Behavioral Science Review* 7 (2): 83–101.
- Press, Gil. 2013. *A Very Short History of Data Science*. Forbes. <https://forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/>.

- Roberts, Steven, Helene Snee, Christine Hine, Yvette Morey, and Hayley Watson, eds. 2016. *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*. London: Springer.
- Seaver, Nick. 2017. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society* 4 (2): 1–12.
- Shapin, Steven, and Simon Schaffer. 1985. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.
- Simondono, Sergio. 2010. *An Introduction to Science and Technology Studies*. Chichester: Wiley-Blackwell.
- Spears, Taylor Clancy. 2014. "Engineering Value, Engineering Risk: What Derivatives Quants Know and What Their Models Do." PhD Thesis, University of Edinburgh.
- Stanley, Liz, and Sue Wise. 1993. *Breaking Out Again: Feminist Ontology and Epistemology*. London: Routledge.
- Turner, Victor. 1969. *The Ritual Process: Structure and Anti-Structure*. Chicago, IL: Aldine.
- Topp, Stephanie M., Marta Schaaf, Veena Sriram, Kerry Scott, Sarah L. Dalglish, Erica Marie Nelson, Rajasulochana SR, et al. 2021. "Power Analysis in Health Policy and Systems Research: A Guide to Research Conceptualisation." *BMJ Global Health* 6 (11): e007268.
- Van Doorn, Niels. 2017. "Platform Labor: On the Gendered and Racialized Exploitation of Low-Income Service Work in the On-Demand' Economy." *Information, Communication & Society* 20 (6): 898–914.
- Van der Sluijs, Jeroen. P. 2002. "A Way Out of the Credibility Crisis of Models Used in Integrated Environmental Assessment." *Futures* 34 (2): 133–146.
- Young, Meg, Michael Katell, and P. M. Krafft. 2019. "Municipal Surveillance Regulation and Algorithmic Accountability." *Big Data & Society* 6 (2): 2053951719868492.
- Ziewitz, Malte. 2016. "Governing Algorithms: Myth, Mess, and Methods." *Science, Technology, & Human Values* 41 (1): 3–16.

About the Author

Daan Kolkman is a computational sociologist and sociologist of computation whose research focuses on the responsible application of algorithms in society. He conducts in-depth qualitative case studies of instances of algorithms used in practice, drawing on the Science and Technology Studies (STS) literature. His research has a strong connection to practice and he works closely with public and private sector organizations to study the impact of algorithms on decision-making.

16. Lessons Learned from The eQuality Project: Privacy and Equality for Youth in Networked Spaces

Valerie Steeves

Abstract

This chapter provides an overview of insights on transdisciplinary research, gained while working on the eQuality Project (EQ), a partnership of academic researchers, civil society and community groups, educators, policymakers, and youth who are working together to explore young people's lived experiences of privacy/surveillance and equality in networked spaces. The article delves into three main areas of the project: research, education/outreach, and policy. It also details working with collaborators, what makes for a successful partnership and what challenges may come up with this type of research.

Keywords: Non-governmental organization; Surveillance; Lived experience; Partnership

Introduction

The eQuality Project (EQ) is an eight-year partnership of academic researchers, civil society and community groups, educators, policymakers, and youth who are working together to explore young people's lived experiences of privacy/surveillance and equality in networked spaces. The project is both interdisciplinary (involving academics from criminology, law, political science, communications, information science, social work, and computer science) and intersectoral (encompassing research, policy work, education, community outreach, and youth engagement). Our partners range in size

from large international groups (e.g., UNICEF Canada) and government departments (e.g., Women and Gender Equality Canada) to grassroots youth organizations (e.g., Canadian Center for Gender and Sexuality Diversity). So when we meet in a room, we encompass a variety of perspectives, agendas, and needs. It is always interesting.

Over the past eight years, we have consciously developed practices that enable this diverse, highly motivated, and sometimes fractious group of people, all of whom have full-time jobs of their own, to collaborate productively on a research project designed to create new knowledge about young people's experiences with networked technologies. In this chapter, I talk about the lessons we have learned along the way, highlighting what has worked and what has not, in the hope that these insights can inform others who are considering a partnership approach.

I start the chapter with a brief overview of the project itself. I then explore the trials and tribulations – and joys – of working with a community of amazingly skilled and dedicated partners. Finally, I conclude with a discussion of specific challenges and benefits brought by this kind of research and consider how this kind of work can inform future research on the datafied society.

The eQuality Project – An Overview

In many ways, EQ grew out of an earlier research collaboration exploring the performance of gender on social media. Our core research team first came together in 2009 because its members shared an interest in mapping what we assumed would be the many new and empowering gender identities that were being expressed in networked spaces. But when we got into the field in 2011, we did not find much empowerment; instead, we were confronted by data that consistently suggested that the commercial design of social media platforms constrained young people's self-representations in ways that magnified the same old stereotypes and set these youth up for peer conflict. At the time, this was new thinking.

As we disseminated our results at a variety of conferences – some academic, some professional – we were repeatedly approached by community organizations and government service providers who were hearing about the same problems from the youth they served. We all agreed on two things: new knowledge was clearly needed; and young people had to be engaged in the process directly, from research through dissemination to policymaking, to ensure that the educational initiatives and policies created by adults for youth would enable them to fully participate in networked life. We

also agreed that we could achieve these goals most effectively by working together, combining the unique perspectives and skills of researchers, youth service providers, educators, community groups, policymakers, and young people.

Accordingly, we combined our efforts to address three interrelated areas: research, education/outreach, and policy.

1. *Research.* In the beginning of the project, we conducted qualitative research to map platform infrastructures and explore how the design of online spaces combines with online social norms to expose young Canadians to discrimination and online harassment. We also used concept mapping and q-sorting to get a better sense of how diverse young people conceptualize both privacy and equality in networked spaces. Since then, we have expanded our efforts in this regard to include qualitative research on young people's understanding of and experiences with artificial intelligence, especially with respect to education. But the most important shift has been in our methods and, as discussed below, we have adjusted our approach due to the influence of our youth partners, who have done the most to actively challenge us, shake us up, and take us in new directions.

2. *Education/Outreach.* Given how many of our partners work directly with youth, we have consistently converted the new knowledge we have created into educational and outreach materials that our partners can use to engage with youth. Some of these materials are more traditional in nature. For example, we have developed a variety of lesson plans that are classroom-ready for teachers across the country. Others are more unusual: we have worked with several young filmmakers to produce short films that give voice to the concerns identified by our research participants and youth partners as important to them. Once again, this shift has been driven by our youth partners.

3. *Policy.* We also hope to contribute to digital media policymaking by fostering public debate and disseminating our new knowledge to policymakers. For example, in addition to traditional policy interventions, the presence of our institutional policy partners has enabled us to hold workshops with policymakers on e-learning in both Alberta and Ontario. Once again, our youth partners have encouraged us to think outside the box, and more recently we have focused on connecting young people directly to policy debates through art and advocacy.

The project received \$2.5 million from the Social Sciences and Humanities Council of Canada (SSHRC). As a condition of the grant, partners must

contribute a minimum of 35 percent of the partnership budget in cash or in-kind contributions. These terms can put pressure on researchers to go where the money is, and in fact both our institutions and SSHRC encouraged us to approach Facebook and/or Google as potential partners. However, early on, our core research team made the decision not to partner with a tech company because we wanted to keep our work as independent as possible. We also decided not to ask our smaller, community-based partners to divert funds from their often stretched budgets to contribute to our collaboration. Remarkably, our institutional partners took up the slack and contributed more than \$1.5 million, knowing that this funding would be distributed across the partnership to supplement the work of our community and civil society partners.

Conceptually, our work is heavily informed by privacy/surveillance theory and by equality theory, framed by a child rights approach that has enabled us to engage directly in the human rights debates about the datafied society. We have also thought long and hard about how our methods would best account for the intersectional nature of young Canadians' lives. Thus we have increasingly sought to center the experiences and needs of marginalized young people in our research and outreach.

Working with Partners

Shifting from traditional research to a partnership approach was not difficult for our team, largely because many of us – especially me, Jane Bailey, and Leslie Shade – had been actively involved in advocating for and with community groups for several years. Whereas tenure and promotion committees often suggest that research and advocacy don't mix well, the deep roots we had laid down throughout our work as advocates meant that we found it easier to build partnerships because we already had strong relationships with educators, civil society groups, and community groups who were active in the area.

These partnerships also gave us a unique perspective on the needs of small, typically under-funded advocacy and community groups. The partnerships I have seen fail in the past were those that privileged academic output such as publications and conference presentations over active engagement with partners. I remember one conversation in particular where advocacy partners were asking for help with legislative interventions – which would have just meant that a student or colleague would dig through the literature and find the data to support the policy they were advocating for – and yet

this assistance was refused because the partnership's academic members noted their tenure and promotion committees wouldn't recognize it as 'valuable work'. I remember, too, working with a small NGO that had accepted an invitation to participate in a grant application only to discover that, once the grant application was accepted, they would never hear from the academic researchers again. I have also seen partnerships where well-funded academics (who are salaried employees of universities) have exerted incredible pressure on their underfunded NGO partners (who are often trying to raise core funding to pay their own salaries) to provide cash and in-kind contributions, which further increased the financial pressure on the NGO without contributing to its mandate in any way.

When we built the EQ partnership, we therefore committed ourselves to: 1) identifying, on an ongoing basis, those areas where our research interests aligned with our partners' interests, not only in knowledge production but also in advocacy, education, and community engagement; and 2) explicitly dedicating a portion of partnership funds to pay for things that can advance our partners' work. This latter commitment is trickier than it sounds, given SSHRC's strict guidelines about what can and cannot be paid to partners. We have paid for student internships, the creation and production of co-branded educational and outreach modules, the co-hosting of public events (e.g., film showings and hackathons), and travel so partners can attend conferences or policy meetings. We have also consciously used our expertise to support partner-driven research, both by volunteering our own time and providing paid student research assistance. These efforts have improved our student training program and have built our partners' capacity to conduct future research on their own.

Mostly, this has worked well. But it is important to remember that partner priorities can shift over time: they will be more or less interested in collaborating with you depending on how well their priorities continue to align with yours. It is also true that different partners will be more involved at different stages of a given project. For example, some of our partners are really interested in research; others are much more interested in using our educational modules in the field. Being flexible and enabling partners to flow in and out of the partnership over time both limits conflict and maximizes output. To help facilitate the best working relationship for all, in addition to our annual in-person meetings and various management committee work, we have always sat down with each partner once a year to see what they are up to and to brainstorm work we could do together. These discussions have also led to a number of invitations for us to participate in partner events, which has been extremely beneficial for our collaborative work.

For example, when I attended a UNICEF workshop with young advocates, I left with a fabulous list of proven strategies that we then used to design our own youth summit. Finding these win-win opportunities is the bread and butter of building a strong partnership.

At the same time, when partner priorities diverge it is really important to be aware that this has happened. We have had differences of opinions about a variety of things, from how to mobilize intersectionality to what it means to 'engage' with youth and whether it is a good idea to frame evaluation from a child-rights approach. When you consider the range of people in the room, it should not be surprising that grassroots youth advocates may have a different perspective than educators or civil servants do, for example. We have been completely blessed by the good will that people have brought to our partnership, and although debate can be fractious it is always respectful. I think one of the things that keeps it that way is that our partners know that they do not have to participate in everything we do. Distance is often very productive, especially for our government partners, who are unable to take sides on many of the debates we engage in because that would interfere with their own mandates. The diversity of opinions in the room has also meant that we have thought much more carefully about our own work as researchers; it is much easier to see the unintended consequences of an approach when you see how different audiences respond to your ideas.

The diversity of needs in the partnership has also opened up a number of new avenues with respect to what we research and how we go about it. Our front-line service providers in particular have been able to highlight areas where new knowledge is desperately needed. For example, one woman had been lobbying for better privacy protection for trans youth to better meet the needs of her clients, but she had been largely dismissed by policymakers because her evidence was anecdotal. Creating research findings was, accordingly, an important way of supporting this kind of evidence-based policy reform.

In addition, having our youth partners review our research protocols has had a significant impact on our methods. Our youth partners have articulated a whole list of things that would be important to young people who participate in research, from certificates for skills training to having us sign off on the volunteer hours they must complete during high school. But most importantly, they have consistently pushed us to take our commitment to youth-driven research to another level. Accordingly, we have expanded our toolkit beyond traditional qualitative and mixed-methods research to include youth participatory action research (YPAR) and arts-based methods. For example, we worked with Dr. Valerie Michaelson to conduct a YPAR

project to help young people critically evaluate the impact of networked technology on their sense of connectedness with themselves, with others, and with nature. We are currently designing a similar YPAR project to explore the mental health needs of trans youth. We have also worked with young artists in Toronto and Ottawa here in Canada and San Juan in Puerto Rico to connect them with artistic mentors and provide them with the materials they need to design and create their own artworks to make their concerns about online surveillance and harassment more visible.

These projects have also enriched our student training program, as our students have been able to make real-world connections between the research enterprise and community action. In turn, our partners have benefitted from student assistance on a wide range of sub-projects, from education and outreach to policy intervention. Students have also been able to create relationships with key individuals working in the sector, which has led to future employment opportunities post-graduation.

Challenges

Although there are many benefits of a partnership approach to research, there are also some consistent challenges. First, managing a long-term collaborative project requires a great deal of time and energy. I urge anyone considering it to build a salary for a full-time project manager into the budget. At the same time, keep in mind that having a project manager is a necessary but insufficient solution to the management problem. Given the kinds of salaries research projects can pay, project managers tend to be young and often (appropriately) use the position as a steppingstone to a permanent job elsewhere. Make sure you have developed systems to track project deliverables in a centralized storage location, so that the entirety of your institutional memory does not leave when your project manager does. A well-thought-out document management system makes the transition much easier.

Second, put the effort in to find out exactly what your university does to support the project, and who does this work. Unfortunately, central services are typically spread across a number of departments, so keep a running list of contacts in research management services, the Office of the Vice-President of Research, and the office of your Faculty Vice-Dean of Research. Build strong relationships with these people; you will definitely be calling them in a panic on at least one occasion to figure out who needs to sign a document that must be submitted to your funder by the end of

business hours. Figuring out who can do what will often be as much of a mystery to them as it will be to you. We once spent eleven months trying to find who in central administration could tell our procurement department that we did not have to tender a job before we could pay a partner for work they had completed for the project.

Third, management only gets more complicated when more than one university is involved. It can take months to figure out how to pay travel expenses or the rental of a conference room, so be proactive and settle those issues in the early stages of planning anything. The same is true of the various institutional research ethics boards that will review your protocols. On our first sub-project, we spent months getting approval from one institution only to have the next institution ask for revisions that were inconsistent with the original certificate, and so on and so on as we applied to the remaining institutions. We learned to approach all five boards in advance and to have a meeting to agree on exactly what would be included in our application so we could submit the same information to each institution at the same time. Then, when we received contradictory requests for revisions, we had another meeting to work out a compromise. For example, on a project working with LGBTQ youth, one university required that we forgo parental consent to protect our participants' privacy in case they had not yet come out to their parents. Another insisted that we have parents not only sign off on their child's participation but that we undertake to inform parents if their child said anything that indicated they might be at risk in some way during the interview. In the end, all five boards agreed to proceed without consent, but it was much easier to resolve the matter because we had already created working relationships, both with them and between them, when we met to discuss the application form.

Fourth, partners are likely to experience staffing changes over the lifetime of the project. Often the partner's institutional commitment to collaboration leaves when your key contact does requiring you to rebuild relationships with partners on an ongoing basis. This takes time, so build it into your plans.

Last, think through the legal issues that may arise when you collaborate. Universities and partners often have conflicting approaches to data protection and copyright issues, for example. This can get tricky, especially when you are trying to co-brand partnership output. Being upfront with everyone early on in the process helps. We have found this is particularly true when we work with youth producers. To avoid conflicts, talking about copyright, ownership, remuneration, and timelines before you start a project limits the possibility that people will go into it with conflicting expectations.

Although all of this requires work, the benefits of the partnership approach are worth the extra effort. It has been a real pleasure to see how the knowledge we create in the academy can be mobilized in the real world to support evidence-based policy and educational initiatives. But the real benefit has been connecting lived experiences of datafication.

References

- Bailey, Jane, and Valerie Steeves (eds). 2015. *eGirls, eCitizens*. Ottawa: University of Ottawa Press.
- Bailey, Jane, Valerie Steeves, Jacquelyn Burkell, Leslie Shade, Priscilla Regan and Rakhi Ruparelia. 2019. "Getting at Equality: Research Methodologies Informed by the Lessons of Intersectionality." *International Journal of Qualitative Methods* 18: 1–13.
- The eQuality Project. "Arts Exchange." Accessed December 31, 2021. <https://www.equalityproject.ca/resources/art-exchange/>.
- The eQuality Project. "Disconnection Challenge." Accessed December 31, 2021. <https://www.equalityproject.ca/resources/disconnectionchallenge/>.
- The eQuality Project. "The eQuality Project." Accessed December 31, 2021. <https://www.equalityproject.ca/>.

About the Author

Valerie Steeves, JD, PhD, is a Full Professor in the Department of Criminology at the University of Ottawa. She is the principal investigator of The eQuality Project, a multi-year partnership of academic researchers, educators, policy analysts and youth funded by the Social Sciences and Humanities Research Council of Canada.

Afterword

Benjamin Peters

Collaborative researchers, so named by this volume, do not seek merely to *study* digital society. We seek to *change* it. The goal has long been to shape our digital society for the better – and indeed that must continue to be our overarching aim.

All of the many ways forward in collaborative research in action involve squaring ourselves with one core recognition: *what we do with data matters*. Consider the chapter on femicide. We can only account for what we can count, so if murders of women do not make it into the official records, there can be no accountability for those murderers, for their murders exist in no societal context. To paraphrase Ferdinand II, “If it is not in the records, it is not in the world.”

Collaborative research is the emerging empirical, practical, and critical response to such a fundamental problem. It calls for many things. By carefully and conscientiously developing and applying data practices with the goal of having societal impact, it will be possible for scholars, researchers, and students to develop vital and rigorous processes of data work, allowing indicators, weights, categories, and orderings of values to be critically challenged, deconstructed, and reconstructed so that they better reflect many different empirical situations.

It is high time that scholars and students of digital societies abandon, at least partially, the dogmatic pretension to theoretical abstraction and welcome instead a welter of conflicting forces as they seek out scientific truthfulness and better patterns through which to collectively understand, rebuild, shape, and govern society. In other words, scholars and students often want to imagine that we can invent powerful frameworks and theories sufficiently complex to solve, unaided, all of society’s complex problems. What this crowd overlooks are the central lessons of this volume: ‘academia is changing’ and ‘do try this at home!’ A theoretical data society confining itself to mere understanding falls short of what can be achieved by the practical sciences that, in understanding the limits of mere understanding, would also attempt to *change* the world.

This shift in perspective is no idle matter. Rather, a focus on what is practically possible comes more clearly into view thanks in part (but only in part) to a radical transformation that has occurred in data analysis over the

last thirty years. While far short of Big Data long-tail mythologies in which no method matters and data are king, it often enough expresses a philosophical fanaticism and the bald, if commonly held, fiction that theory, understood as a normative framework, exists independently of the available data on a given project. It is rather the case that data about a problem gathered *en masse* may, if accompanied by careful reflection on the origins, biases, and consequences of that data, help the researcher construct a grounded theory that can drive practical problem-solving.

Setting aside philosophical fanaticism, wherever it may exist, may increase our ability for discerning observation and mutual sympathy, as well as our capacity to discard rose-colored and dystopian perspectives alike that can be found in the cyclicity of the criti-hype digital discourse. Instead our hope as mediated creatures is not to reject in ignorance or embrace in exhaustion the warped, yet deeply embedded digital realities we inhabit; rather, our hope lies in observing, taking account of, measuring, and pragmatically sorting through the case for social orders and common goods that are far greater than the solitary observer or the selfish self – and doing so together.

Most of the contents of this volume, as well as most of what constitutes a collaborative framework worth building on – good conversations, cooperative research, service with an eye to something larger than personal reward, and community with a goal greater than institutional profit – all have taken root and sprout in what are now natively digital and interpersonal soils of what it means to be a good digital citizen: participation, in good faith, in projects based on network forums, unfolding in neighborhoods and communities, public spaces and parks, classrooms at community centers and among congregations, involving the payment of local and regional taxes, advocacy for policy reform that benefits those most in need, and the democratic rituals that transpire at polling places. Most of the action, and therefore most of the research that must follow, exists in the space of collaboration, namely the digital equivalents of middle-school gyms, neighborhood councils, laboratories, city councils, and state assemblies.

The editors, in their introduction, point to “the elephant in the room” – and what an elephant it is! The myths of the selfish solitary self – hardy perennials – are embedded in the soil of much of the digital discourse. Collaboration offers a healthy antidote to such brutish braying and the stampedes of selfishness, which can otherwise be hard to chase out. Consider, for example, how, at least in the United States, the field of computer programming education often attracts students via a bait-and-switch: the young student programmer is lured into the technical side of digital society research by the promise of becoming a universal analyst, free to navigate

the world of digital knowledge. Promised the powers of veritable mage, they are instead trained and marketed to organizations as increasingly narrow and insider specialists with hyper-specific skills: content moderators, coders, and human trainers of algorithms. The idea of the user as a coherent self supports another pervasive myth, that of meritocracy, fostering complacency towards (a lack of) diversity and fairness in the workplace. However, measuring the purported merit of the specialist self often takes into account only a very narrow sum of cumulative advantages accruing to a person over a given lifetime, and it ignores how digital society reflects practical bottlenecks and biases in the hiring pipeline. For this reason, public coding education movements often fail to recognize that, if they successfully achieve their goals, they will launch a future wave of coders with technical programming talents into the job market without addressing underlying industry biases in their hiring practices. The problems besetting digital society today are far deeper than can be remedied by simply devoting equitable attention to sex, gender, race, ethnicity, class, culture, community, language, platform, infrastructure, nationality, and other identity categories, although intersectional analyses of these variables often help institutions make vital corrections in areas where they have fallen short. Between these often sensitive categories lies much that is worthy of articulation but that nonetheless slips between the cracks of trivially scalable technical analysis.

A proper collaborative approach to data practice would ease many of these pressures, and such an approach should extend beyond technical-only teams. Consider another analytical paradox faced by teams exclusively made up of data engineers, analysts, and technical scientists: they possess precisely the tools to understand how subpopulations within a population are *not* similar to one another, and yet at the same time they are institutionalized and incentivized to solve problems affecting more than a single population. Moreover, even the most accurate data representation of a subpopulation may alone suffice to *describe* a problem to be addressed by future action, but description alone, without the necessary but missing context, is not actionable: reality, as Kant recognized, cannot be reduced to only that which can be represented. In other words, the typical approach taken by a solitary engineer (or, really, any single analytical approach) to address digital societal issues – represent the issue at hand better – may be necessary but will never be sufficient. Instead, this volume calls for the rerouting of our attention: away from the universal, the global, and the trivially scalable information system problems and towards the immediate, local, and regional contexts of lived experience, inviting listening and learning across diverse regions of data-informed societal collaboration. We are called, too, to revel in how

such a transformation toward transdisciplinary collaboration involves not a simplification, but in fact a flowering of dimension, color, and complexity.

This volume cries out with two cheers for the complexity of the case study. The robust capacity of local and regional situations to surprise us should suffice to check the analytical instinct to infinitely generalize and trivially systematize. Consider how often our case studies about the good life in datafied societies surprise us. For example, a reporting requirement meant to protect victims of abuse in one situation may also expose these people to worse risks in another, as has been shown in these pages. Similarly, common practice may suggest that the safest way to secure one's privacy is to trust one particular institution, but what happens when its data are breached? Sometimes the fate of prognosticators surprises in the other direction: Thomas Hobbes, remembered for declaring life to be "solitary, poor, nasty, brutish, and short", passed away in his bed surrounded by friends and family, having almost reached his nineties, more than twice the average life expectancy of his era. In both extremes, the greatest cautions in life demand to be assessed case by case. Perhaps for this reason, the laws of many countries recognize the novel and profound reality of the case study, for the case study disrupts the utopian dreamers and the dystopian catastrophists alike. Even as regional trends can and must continue, the stream of history will go on rippling and refreshing in somewhat, but never fully, predictable ways.

The contributors to this volume have advanced precisely such a practical approach to studies of digital societies seeking to make a difference. Their case studies have been selected and framed, in the best tradition of the pragmatists, with a clear sense of what will most likely work best. Social, civic, and political questions appear in a large share of these pages. As a whole, the volume models the shedding of any traditional theoretical orthodoxy that would root the analyst either in a highly narrow empirical case study that cannot see outside itself or in the impenetrable fog of a general theory that cannot distinguish itself from fiction, and it refuses to adopt yet another orthodoxy. It may not be too much of a stretch to assert that, insofar as this volume has or even needs to admit a theory of truth, it is perhaps best to say that it follows John Dewey's notion of "instrumentalism": the best or at least a first principle applicable to collaborative data societal study is that each study take up the best practices for its case, and that its processes be understood to be temporal, evolutionary, and oriented towards making a practical difference in the lives of all those the study touches. Under what conditions would a given study be most likely to surprise its

creators by correcting a claim as false that was previously held to be true as well as by, in the process of being corrected, changing the situation at hand to make it better? Those pre-scientific conditions and normative evaluations, co-constituting one another and never fully separable in the process we call data science (or once we admit the social construction of data and scientific practice, we may call simply “science”), provide the best (or most instrumental) meaning for that claim.

In gathering studies that pair critical practical studies of infrastructure and ethnography, data regimes and data practices, partnerships and participatory research across civic society, policy and politicized research, and most powerfully a long line of particular case studies whose relevance burst beyond their own bounds, this volume seeks to be an antidote for, and an alternative to, the popular approach to averaging out the best data practices as if they were subject to a central limit theorem: if the last time was too large, go smaller; if too small, go larger. No, none of that crude instrumentalization please! Rather, via applied theorizations as well as case study analyses, this volume models a critical approach that can foster better instrumentalist data practices. The volume does not aim to put forth absolutely true or false judgements; rather, it seeks to model a higher art of data-informed inquiry in which there must be continuous mutual adjustments and adaptations between the environment, object, and subject under analysis. This is not to diminish the meaning or significance of what can be accomplished through data-informed societal study. Neither does it cede ground to critics who would be too quick to accuse such an approach of being ‘under-complex’: rather, in precisely the philosophical move that critics of practical data science often miss, this volume, in its practice, slightly redefines and elevates the very definitions of *meaning* and *significance*. Motivated by no mere *gesellschaftsgestalterisch* impetus, collaborative research also seeks an epistemic impetus: to paraphrase Charles Sanders Peirce, meaning, paid out in the soil of facts that change lives, is precisely what matters. We should seek to understand the meaning and significance of data research precisely via the full sum of their practical effects. What else could a better data-informed society look like except one that, for once, understood its meaning and significance to be precisely that which makes a difference to those in all cases at hand, and thus the task of data-informed societal inquiry is to make a difference that matters? Without predetermined advocacy, this volume models a tradition as impactful as it is self-reflective, mounting an evolutionary enquiry that deserves attention, imitation, and its full flowering in the consequences its readers will realize with effort and over time.

There's nothing simple about such a call, and those who would see it as under-complex have fundamentally misunderstood the task at hand. It is not that complexity lies in framing one's work with theory but rather that practice lies in complexity. Practice written down is sufficiently complex to fill libraries – and every other institution. To transform data analytic research from the erudite, individual, and headily abstract into practical, engaged, and collaborative analysis in action requires significant, even enormous, amounts of effort and time, as well as complex investments, intellects, and infrastructure. Every dataset arrives in the hands of the researcher pre-cooked by different situations, categories, and contexts that are social, economic, political, and institutional. Thus to understand how to collaboratively engage that data into practices that improve the world requires simultaneously a search to understand the world itself – or at least the particular world that generated that data – and reflection on the practices through which it might be engaged. In a very specific way, by seeking to grasp the phenomenon at hand, through its different shades and situatedness, the contingencies and interdependencies in its local, social, economic, political, and institutional being, we gain not just a practical understanding of the phenomenon in question but a general understanding of its place in the world.

This epistemic commitment to understand data phenomena practically also ensures that the methodologically complex toolset that data practitioners bring to their work – digital methods, qualitative methods, quantitative methods from different industries – will be understood through a thoroughgoing commitment to interdisciplinarity and transdisciplinarity. A practical response to a multipart question such as “When does summer end in Utrecht, and how and why has that answer changed over time?” may require connecting academic scholars such as historians of climate, philosophers of seasonality and quantitative methods, and cultural anthropologists with various weather, climate, and epistemologically sensitive data scientists at local, municipal, regional, national, and transnational levels. How much more complex and collaborative, then, should, and must, even more complex questions become? I'm thinking here of questions about our changing climate, political economies, societies, cultures, and lives. Surely the baseline for living intelligently in a datafied society compels and indeed requires academia to collaborate with societal sectors and to engage with external partners, along with allocating funding, translating various ontologies and professional practices, building and sustaining social networks of researchers and practitioners, supporting publications in academic, professional, and public venues, and just developing, in general,

novel forms of knowledge and insight for, and with, the student and the professional alike. It is the single ambition of the editors of this volume to make this often invisible, unacknowledged, and unrewarded labor visible and, moreover, imitable, praiseworthy, and promotable. With this volume in hand, may universities, societal organizations, industries, and governments near and far recognize, reward, facilitate, adopt, and model for others what it means, and with effort and time must still come to mean, to conduct collaborative research in a datafied society.

About the Author

Benjamin Peters is a media scholar and author and editor of several books. 2022–2023 Senior Fellow at the Cultures of Research KHK at RWTH, Aachen, Germany, he is the Hazel Rogers Associate Professor of Media Studies at the University of Tulsa, Oklahoma, USA and affiliated fellow at the Information Society Project at Yale Law School, Connecticut, USA.

Bibliography

- Abolafia, Mitchel Y. 1999. "Markets as Cultures: An Ethnographic Approach." *Sociological Review* 46 (1_suppl): 69–85.
- Acey, Camille E, Siko Bouterse, Sucheta Ghoshal, Amanda Menking, Anasuya Sengupta, and Adele G Vrana. 2021. "Decolonizing the Internet by Decolonizing Ourselves: Challenging Epistemic Injustice through Feminist Practice." *Global Perspectives* 2 (1): np.
- Acker, Amelia. 2015. "Toward a Hermeneutics of Data." *IEEE Annals of the History of Computing* 37 (3): 70–75. <https://doi.org/10.1109/MAHC.2015.68>.
- Adler, Patricia A., and Peter Adler. 1987. *Membership Roles in Field Research*. Vol. 6. Newbury Park: Sage.
- Afinogenov, Greg. 2018. "Orthodoxxed! On 'Sokal Squared.'" *n+1* (blog). October 4, 2018. <https://www.nplusonemag.com/online-only/online-only/orthodoxxed/>.
- Akhter, Rawshon. 2020. "Research That Hurts – On the Practice of Care in Fieldwork." *FemLab.Co* (blog). June 26, 2020. <https://femlab.co/2020/06/26/research-that-hurts-on-the-practice-of-care-in-fieldwork/>.
- Allen, Chris. 2011. "Against Dialogue: Why Being Critical Means Taking Sides Rather Than Learning How to Play the 'Policy Research' Game." *Dialogues in Human Geography* 1 (2): 223–227.
- Anderson, Claire, and Paul Bissell. 2004. "Using Semi Covert Research to Evaluate an Emergency Hormonal Contraception Service." *Pharmacy World and Science* 26 (2): 102–106.
- Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. "Machine Bias." ProPublica, March 23, 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Animiiki. 2019. "Decolonizing Digital: Contextualizing Indigenous Data Sovereignty." <https://animikii.com/news/decolonizing-digital-contextualizing-indigenous-data-sovereignty>
- Anwar, Mohammad Amir, and Mark Graham. 2021. "Between a Rock and a Hard Place: Freedom, Flexibility, Precarity and Vulnerability in the Gig Economy in Africa." *Competition & Change* 25 (2): 237–258. <https://doi.org/10.1177/1024529420914473>.
- Bacchi, Carol. 2009. *Analysing Policy*. Frenches Forrest, NSW: Pearson Higher Education AU.
- Bailey, Jane, and Valerie Steeves (eds). 2015. *eGirls, eCitizens*. Ottawa: University of Ottawa Press.
- Bailey, Jane, Valerie Steeves, Jacquelyn Burkell, Leslie Shade, Priscilla Regan and Rakhi Ruparelia. 2019. "Getting at Equality: Research Methodologies Informed by the Lessons of Intersectionality." *International Journal of Qualitative Methods* 18: 1–13.

- Bakker, Jeroen, Daphne Cornelisse, Sahra Mohamed, Mirko Tobias Schäfer and Joris Veerbeek. 2021. "Van Scherm naar Straat. Hoe Sociale Media-Conversations Protest op Straat Mobiliseren." Research report. Utrecht Data School. Utrecht University.
- Bansal, Pallavi. 2021. "Is Feminist Design a Solution to Platform Workers' Problems?" *FemLab.Co* (blog). February 20, 2021. <https://femlab.co/2021/02/20/is-feminist-design-a-solution-to-platform-workers-problems/>.
- Banting, Keith, and Debra Thompson. 2021. "The Puzzling Persistence of Racial Inequality in Canada." *Canadian Journal of Political Science/Revue canadienne de science politique* 54 (4): 870–891.
- Bardzell, Shaowen. 2010. "Feminist HCI: Taking Stock and Outlining an Agenda for Design." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1301–1310. CHI '10. New York, NY: ACM. <https://doi.org/10.1145/1753326.1753521>.
- Barton, Allen H. 1979. "Paul Lazarsfeld and Applied Social Research: Invention of the University Applied Social Research Institute." *Social Science History* 3 (3–4): 4–44.
- Bastow, Simon, Patrick Dunleavy, and Jane Tinkler. 2014. *The Impact of the Social Sciences: How Academics and Their Research Make a Difference*. London: Sage.
- Bates, Tony. 2010. "New Challenges for Universities: Why They Must Change." In: *Changing Cultures in Higher Education: Moving Ahead to Future Learning*, edited by Ulf-Daniel Ehlers and Dirk Schneckenberg, 15–25. Berlin: Springer.
- Beck, Ulrich. 1992. *The Risk Society: Towards a New Modernity*. London: Sage.
- Bernal Sarmiento, Camilo, Miguel Lorente Acosta, Françoise Roth, and Margarita Zambrano. 2014. *Modelo de protocolo latinoamericano de investigación de las muertes violentas de mujeres por razones de género (femicidio/feminicidio)*. ONU Mujeres; OACNUDH.
- Berry, David M. 2012. *Understanding Digital Humanities*. Houndmills, Basingstoke and Hampshire: Palgrave Macmillan.
- Berry, David M., Erik Borra, Anne Helmond, Jean-Christophe Plantin, and Jill Walker Rettberg. 2015. "The Data Sprint Approach: Exploring the Field of Digital Humanities through Amazon's Application Programming Interface." *Digital Humanities Quarterly* 9 (3).
- Bertolini, Alessio, Adam Badger, Fabian Ferrari, Mark Graham and Funda Ustek-Spilda. Forthcoming. "Reversing the Race to the Bottom in the Platform Economy: The Fairwork Approach." In *Handbook of Labour Geography*, edited by Andrew Herod.
- Beunza, Daniel, and David Stark. 2012. "From Dissonance to Resonance: Cognitive Interdependence in Quantitative Finance." *Economy and Society* 41 (3): 383–417.

- Beunza, Daniel, and Raghu Garud. 2007. "Calculators, Lemmings or Frame-makers? The Intermediary Role of Securities Analysts." *Sociological Review* 55 (2_suppl): 13–39.
- Birkbak, Andreas, Morten Krogh Petersen, and Torben Elgaard Jensen. 2015. "Critical Proximity as a Methodological Move in Techno-Anthropology." *Techné: Research in Philosophy and Technology* 19 (2): 266–290. <https://doi.org/10.5840/techne201591138>.
- Biruk, Cal. 2018. *Cooking Data: Culture and Politics in an African Research World*. Illustrated edition. Durham, NC: Duke University Press.
- Black, Max. 1962. *Models and Metaphors: Studies in Language and Philosophy*. Ithaca, NY: Cornell University Press.
- Blume, Stuart. 1985. "After the Darkest Hour: Integrity and Engagement in the Development of University Research." In *The University Research System: The Public Policies of the Home of Scientists*, edited by B. Wittrock and A. Elzinga, 139–163. Stockholm, Sweden: Almqvist & Wiksell International.
- Bogner, Alexander. 2021. *Die Epistemisierung des Politischen. Wie die Macht des Wissens die Demokratie gefährdet*. Stuttgart: Reclam.
- Böhme, Gernot, and Nico Stehr. 1986. *The Knowledge Society: The Growing Impact of Scientific Knowledge on Social Relations*. Vol. 10. Dordrecht: Springer Science & Business Media.
- Boiten, Max, Maica Hopstake, Marjolein Krijgsman, Roelof Muis, Peter Mullaart, Sander Prins, Mirko Tobias Schäfer, Joris Veerbeek. 2020. "Online Bedreigingen in het Lokaal Bestuur." In *Aard en Omvang van Dader- en Slachtofferschap van Cyber- en Gedigitaliseerde Criminaliteit in Nederland*, edited by M.G.C.J. Beerthuizen, T. Sipma, and A. M. van der Laan, vol. Cahier 2020–15. 46–58. The Hague: WODC.
- Bok, Derek. 2003. *Universities in the Marketplace: The Commercialization of Higher Education*. Princeton, NJ: Princeton University Press.
- Borden, Kyla. 2017. *Raising the Open Government Bar: Changes to Canada's Access to Information Act*. Retrieved from: <https://www.opengovpartnership.org/stories/raising-the-open-government-bar-changes-to-canadas-access-to-information-act/>.
- Bounegru, Liliana, and Jonathan Gray, eds. 2021. *The Data Journalism Handbook: Towards a Critical Data Practice*. Amsterdam: Amsterdam University Press. <https://www.aup.nl/en/book/9789462989511/the-data-journalism-handbook>.
- Bourgois, Philippe. 2007. "Confronting the Ethics of Ethnography: Lessons From Fieldwork in Central America." In *Ethnographic fieldwork: An anthropological reader*, edited by A.C.G.M. Robben & J.A. Sluka, 288–297. Hoboken, New Jersey: Wiley-Blackwell.

- Bowker, Geoffrey C., and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- boyd, danah, and Kate Crawford. 2011. "Six Provocations for Big Data." *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*.
- boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon." *Information, Communication & Society* 15 (5): 662–679.
- Branderhorst, Martiene. 2020. "Digitaal Leiderschap: Verkenning van de Veranderende Rol van Gemeentesecretarissen in de Informatiesamenleving." *Bestuurswetenschappen* 74 (4): 26–43. <https://doi.org/10.5553/Bw/016571942020074004003>.
- Brandt, Reinhard. 2011. *Wozu noch Universitäten?: Ein Essay*. Hamburg: Felix Meiner Verlag.
- Brannick, Teresa, and David Coghlan. 2007. "In Defense of Being "Native": The Case for Insider Academic Research." *Organizational research methods* 10 (1): 59–74.
- Brevini, Benedetta. 2021. "Creating the Technological Saviour: Discourses on AI in Europe and the Legitimation of Super Capitalism." In *AI for Everyone?: Critical Perspectives*, edited by Pieter Verdegem, 145–160. London: University of Westminster Press.
- Broussard, Meredith. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.
- Brown, Tim. 2009. *Change by Design*. New York, NY: Harper Business.
- Bruno, Isabelle, Emmanuel Didier, and Tommaso Vitale. 2014. "Statactivism: Forms of Action Between Disclosure and Affirmation." *Partecipazione e Conflitto – Participation and Conflict* 7 (2): 198–220.
- Bruno, Isabelle, Florence Jany-Catrice, and Béatrice Touchelay (eds). 2016. *The Social Sciences of Quantification: From Politics of Large Numbers to Target-Driven Policies*. New York, NY: Springer.
- Bueno-Hansen, Pascha. 2010. "Femicidio. Making the Most of an 'Empowered Term'." In *Terrorizing Women: Femicide in the Americas*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, translated by Graciela Trevisan, 290–311. Durham: Duke University Press.
- Buolamwini, Joy. 2023. *Unmasking AI: My Mission to Protect What Is Human in a World of Machines*. New York, NY: Random House.
- Burawoy, Micheal. 2005. "For Public Sociology." *American Sociological Review* 70: 4–28.
- Buteau, Sharon, Aurelie Larquemin and Jyoti Prasad Mukhopadhyay and Sharon. 2016. "Open Data and Evidence-Based Socio-Economic Policy Research in India: An Overview." *The Journal of Community Informatics* 12 (2): 120–147.

- CA-ICT, Nederland ICT, and CIO Platform Nederland. 2017. "Arbeidsmarktonderzoek Topsectoren 2017." <https://www.caict.nl/arbeidsmarktonderzoek-topsectoren/>.
- Calvey, David. 2008. "The Art and Politics of Covert Research: Doing 'Situated Ethics' in the Field." *Sociology* 42 (5): 905–918.
- Canclini, Nestor Garcia. 2008. "Imaginários Culturais da Cidade: Conhecimento/Espectáculo/Desconhecimento." In *A Cultura pela Cidade*, organized by Teixeira Coelho, 15–32. São Paulo: Ed. Iluminuras: Itaú Cultural.
- Canney, Julia. 2021. "Establishing a National Femicide Watch in the United States." *The Hill* (blog). 8 May 2021. <https://thehill.com/opinion/civil-rights/566544-establishing-a-national-femicide-watch-in-the-united-states>.
- Cant, Callum. 2019. *Riding for Deliveroo: Resistance in the New Economy*. Cambridge: Polity Press.
- Cardoso, Ana Cláudia Duarte. 2007. *O Espaço Alternativo: Vida e Forma Urbana nas Baixadas de Belém*. Belém: Editora Universitária da UFFa.
- Caretta, Martina Angela, and Yvonne Riaño. 2016. "Feminist Participatory Methodologies in Geography: Creating Spaces of Inclusion." *Qualitative Research* 16 (3): 258–266. <https://doi.org/10.1177/1468794116629575>.
- Carey, James. 1978. "The Ambiguity of Policy Research." *Journal of Communication* 28 (2): 114–119.
- Castells, Manuel. 2017. *Networks of Outrage and Hope: Social Movements in the Internet Age*. Cambridge: Polity Press.
- CEPAL. 2018. *Mujeres afrodescendientes en América Latina y el Caribe. Deudas de igualdad*. Santiago: CEPAL.
- CFOJA. n.d. "Femicide and the Law | Femicide in Canada". Canadian Femicide Observatory for Justice and Accountability. Accessed November 3, 2021. <https://www.femicideinCanada.ca/courts>.
- Chalabi, Mona, and Jonathan Gray. 2021. "Sketching with Data." In *The Data Journalism Handbook: Towards a Critical Data Practice*, edited by Liliana Bounegru and Jonathan Gray, 174–181. Amsterdam: Amsterdam University Press. <https://www.aup.nl/en/book/9789462989511/the-data-journalism-handbook>.
- Chen, Mel. 2012. *Animacies: Biopolitics, Racial Mattering, and Queer Affect*. Durham, NC: Duke University Press.
- Chenou, Jean-Marie, and Carolina Cepeda-Másmela. 2019. "#NiUnaMenos: Data Activism From the Global South". *Television & New Media* 20 (4): 396–411. <https://doi.org/10.1177/1527476419828995>.
- Christin, Angèle. 2020. "The Ethnographer and the Algorithm: Beyond the Black Box." *Theory and Society* 49 (5): 897–918.
- Citizenship and Immigration Canada (CIC). 2013. *Local Immigration Partnership Handbook*. <http://p2pcanada.ca/wp-content/uploads/2013/02/CIC-Local-Immigration-Partnerships-Handbook.pdf>.

- Cizek, Katerina and William Uricchio. 2022. *Collective Wisdom: Co-Creating Media for Equity and Justice*. Cambridge, MA: MIT Press.
- CLEAR. 2021. "CLEAR Lab Book." CLEAR Lab Book (blog). 2021. <https://civiclaboratory.nl/clear-lab-book/>.
- Cobham, Alex, Jonathan Gray, and Richard Murphy. 2017. "What Do They Pay? Towards a Public Database to Account for the Economic Activities and Tax Contributions of Multinational Corporations." January 2017. CIYPERC Working Paper Series. London: City University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3049857.
- Coleman, E. Gabriella. 2013. *Coding Freedom: The Ethics and Aesthetics of Hacking*. Princeton, NJ: Princeton University Press.
- Collini, Stefan. 2012. *What Are Universities For?* Penguin UK.
- Collins, Harry, and Robert Evans. 2008. *Rethinking Expertise*. Chicago, IL: University of Chicago Press.
- Collins, Harry. 2004. "Interactional Expertise as a Third Kind of Knowledge." *Phenomenology and the Cognitive Sciences* 3 (2): 125–143.
- Coté, Mark, and Jennifer Pybus. 2016. "Simondon on Datafication. A Techno-Cultural Method." *Digital Culture & Society* 2 (2): 75–92. <https://doi.org/10.14361/dcs-2016-0206>.
- Couldry, Nick, Aristeia Fotopoulou, and Luke Dickens. 2016. "Real Social Analytics: A Contribution Towards a Phenomenology of a Digital World." *British Journal of Sociology* 67 (1): 118–137.
- Couldry, Nick. 2014. "Inaugural: A Necessary Disenchantment: Myth, Agency and Injustice in a Digital World." *Sociological Review* 62 (4): 880–897. <https://doi.org/10.1111/1467-954X.12158>.
- Couldry, Nick. 2017. "The Myth of Big Data." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 235–240. Amsterdam: Amsterdam University Press.
- Couldry, Nick. 2019. "Afterword." In *Citizen Media and Practice: Currents, Connections, Challenges*, edited by Hilde Stephansen and Emiliano Treré. London: Routledge. 181–189.
- Crawford, Kate. 2021. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.
- Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica: Journal of the Econometric Society*: 1431–1451.
- Creswell, John W., Cheryl N. Poth. 2017. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: SAGE.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.

- D'Ignazio, Catherine, Helena Suárez Val, Silvana Fumega, Harini Suresh, Isadora Araujo Cruxên, Wonyoung So, María de los Ángeles Martínez, and Mariel García-Montes. 2020. 'Femicide & Machine Learning: Detecting Gender-Based Violence to Strengthen Civil Sector Activism'. In *Mechanism Design for Social Good (MD4SG '20)*. online. <https://youtu.be/1EIDZ443wR4>.
- D'Ignazio, Catherine. 2023. "Human-Centered Computing and Femicide Counterdata Science." In *The Routledge International Handbook of Femicide and Femicide*, edited by Myrna Dawson and Saide Mobayed Vega, 529–541. New York, NY: Routledge.
- Dalton, Craig M., Linnet Taylor, and Jim Thatcher. 2016. "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* 3 (1): 2053951716648346. <https://doi.org/10.1177/2053951716648346>.
- Dalton, Craig, and Jim Thatcher. 2014. "What Does a Critical Data Studies Look Like and Why Do We Care?" *Society & Space blog*, May 19, 2014. Accessed December 1, 2020. <http://www.societyandspace.org/articles/what-does-a-critical-data-studies-look-like-and-why-do-we-care>.
- Das, Djurre, Patricia Faasse, Bart Karstens, and Paul Diederer. 2020. *Raad Weten met Digitalisering: Hoe de Gemeenteraad Kan Sturen op de Maatschappelijke Impact van Digitale Technologie*. The Hague: Rathenau Instiute.
- Davenport, Thomas H. 2020. "Beyond Unicorns: Educating, Classifying, and Certifying Business Data Scientists." *Harvard Data Science Review* 2 (2). <https://doi.org/10.1162/99608f92.55546b4a>.
- Davenport, Thomas H., and D. J. Patil. 2012. "Data Scientist: The Sexiest Job of the 21st Century." *Harvard Business Review* 90 (5): 70–76.
- Davis, Angela Y. 2003. *Are Prisons Obsolete?* Open Media Book. New York, NY: Seven Stories.
- de Stefano, Valerio. 2015. *The Rise of the 'Just-in-Time Workforce': On-Demand Work, Crowd Work and Labour Protection in the 'Gig-Economy'*. ILO Conditions of Work and Employment Series No.71. Geneva: ILO.
- Dekker, C. M., Groenendijk, A., Sliggers, C. J., and Verboom, G. K. 1990. "Quality Criteria for Models to Calculate Air Pollution." *Lucht (Air)* 90. Leidschendam: Ministry of Housing, Physical Planning and Environment.
- Delmar, Rosalind. [2001] 2018. "What Is Feminism?" In *Theorizing Feminism: Parallel Trends in the Humanities and Social Sciences*, edited by Anne C. Herrmann and Abigail J. Stewart, 5–28. London: Routledge.
- Denton, Emily, Alex Hanna, Razvan Amironesei, Andrew Smart, Hilary Nicole, and Morgan Klaus Scheuerman. 2020. "Bringing the People Back In: Contesting Benchmark Machine Learning Datasets." *Proceedings of ICML Workshop on Participatory Approaches to Machine Learning*. ArXiv:2007.07399 [Cs], July. <http://arxiv.org/abs/2007.07399>.

- Department for Transport 2020. "Report of the Laidlaw Inquiry." Available online at: <https://www.gov.uk/government/publications/report-of-the-laidlaw-inquiry>. Last accessed on January 28, 2024.
- Department of Premier and Cabinet. 2017. *Family Violence Outcomes Framework and Indicators*. Melbourne: Victorian Government. <https://www.vic.gov.au/family-violence-outcomes-framework>.
- Desrosières, Alain. 2002. *The Politics of Large Numbers: A History of Statistical Reasoning*, translated by Camille Naish. Cambridge, MA: Harvard University Press.
- Deus, Alicia, and Diana Gonzáles. 2018. "Analysis of Femicide/Feminicide Legislation in Latin America and The Caribbean and a Proposal for a Model LAW." UN Women – United Nations Entity for Gender Equality and the Empowerment of Women. <https://lac.unwomen.org/en/digiteca/publicaciones/2018/12/analisis-legislacion-feminicidio-femicidio-modelo-de-ley>.
- Dijk, José van, Thomas Poell, and Martijn de Waal. 2016. *De Platform Samenleving: Strijd Om Publieke Waarden in Een Online Wereld*. Amsterdam: Amsterdam University Press.
- Dijk, José van, Thomas Poell, and Martijn De Waal. 2018. *The Platform Society: Public Values in a Connective World*. New York, NY: Oxford University Press.
- Dooley, Lawrence, and David Kirk. 2007. "University-industry Collaboration: Grafting the Entrepreneurial Paradigm onto Academic Structures." *European Journal of Innovation Management* 10 (3): 316–332. <https://doi.org/10.1108/14601060710776734>.
- Dragiewicz, Molly and Jean Burgess. 2016. "Domestic Violence on #qanda: The 'Man' Question in Live Twitter Discussion on the Australian Broadcasting Corporation's Q&A." *Canadian Journal of Women and the Law*, 28 (1): 211–229. <https://doi.org/10.3138/cjwl.28.1.211>.
- Dubet, Francois. 1996. "¿Ocaso de la idea de sociedad?" *Revista de Sociología* (10): 7–23. <https://doi.org/10.5354/0719-529X.1996.27653>.
- Dumit, Joseph. 2018. "The Other Ninety per Cent: Thinking with Data Science, Creating Data Studies." In *Ethnography for a Data Saturated World*, edited by Hannah Knox and Dawn Nafus, 252–274. Manchester: Manchester University Press. <https://www.manchesterhive.com/view/9781526127600/9781526127600.00020.xml>.
- DuMonthier, Asha, Chandra Childers, and Jessica Milli. 2017. "The Status of Black Women in the United States." *IWPR* (blog). 7 June 2017. <https://iwpr.org/iwpr-issues/race-ethnicity-gender-and-economy/the-status-of-black-women-in-the-united-states/>.
- Dutta, Mohan, Srividya Ramasubramanian, Mereana Barrett, Christine Elers, Devina Sarwatay, Preeti Raghunath, Satveer Kaur, Debalina Dutta, Pooja Jayan, and Mahbubur Rahman. 2021. "Decolonizing Open Science: Southern Interventions." *Journal of Communication* 71 (5): 803–826.

- Dwivedi, Yogesh K., et al. 2024. "Real impact: Challenges and Opportunities in Bridging the Gap Between Research and Practice—Making a difference in Industry, Policy, and Society." *International Journal of Information Management* 78:102750.
- Dwyer, Sonya Corbin, and Jennifer L. Buckle. 2009. "The Space Between: On Being an Insider-Outsider in Qualitative Research." *International Journal of Qualitative Methods* 8 (1): 54–63.
- Economist Intelligence Unit. 2021. "Democracy Index 2020: In Sickness and in Health?" <https://www.eiu.com/n/campaigns/democracy-index-2020/>.
- Edwards, Paul N., Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker, and Christine L. Borgman. 2011. "Science Friction: Data, Metadata, and Collaboration." *Social Studies of Science* 41 (5): 667–90. <https://doi.org/10.1177/0306312711413314>.
- El-Sammak, Katherine. 2016. "Measuring Open Government: Critically Reviewing OGP's Footprint." November 17, 2016. <https://www.opengovpartnership.org/stories/measuring-open-government-critically-reviewing-ogps-footprint/>.
- Espeland, Wendy Nelson, and Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology / Archives Européennes de Sociologie* 49 (3): 401–436.
- Ettlinger, Krista M. 2021. *Transdisciplinary Social Learning and Transdisciplinary Capacity: Responding to the Challenge of Transdisciplinary Knowledge Integration in the Public Sector*. Master's thesis, Utrecht University School of Governance.
- Etzkowitz, Henry. 2003. "Research Groups as 'Quasi-Firms': The Invention of the Entrepreneurial University." *Research Policy* 32 (1): 109–121.
- Etzkowitz, Henry. 2014. "Making a Humanities Town: Knowledge-Infused Clusters, Civic Entrepreneurship and Civil Society in Local Innovation Systems." *Triple Helix* 2 (December): 1. <https://doi.org/10.1186/s40604-014-0012-z>.
- Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St Martin's Press.
- Fairwork. 2024. Location-based Platform Work Principles. Accessed February 14, 2024. <https://fair.work/en/fw/principles/fairwork-principles-location-based-work/>.
- Fals Borda, Orlando. 1987. "The Application of Participatory Action-Research in Latin America." *International Sociology* 2 (4): 329–347.
- Farmer, Jane, Anthony McCosker, Kath Albury, Amir Aryani. 2023. *Data for Social Good: Non-Profit Sector Data Projects*. London: Springer Nature.
- Ferrari, Fabian, and Mark Graham. 2021. "Fissures in Algorithmic Power: Platforms, Code, and Contestation." *Cultural Studies* 35 (4-5): 814–832. <https://doi.org/10.1080/09502386.2021.1895250>.
- Fischer, Frank. 2005. "Beyond Empiricism: Policy Inquiry in Post-positivist Perspective." *Policy Studies Journal* 26 (1): 129–146.
- Fischer, Frank. 2009. *Democracy and Expertise: Reorienting Policy Inquiry*. New York, NY: Oxford University Press.

- Floridi, Luciano, and Mariarosaria Taddeo. 2016. "What is Data Ethics?" *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2083): 20160360.
- Fox, Jesse, Katy E. Pearce, Adrienne L. Massanari, Julius Matthew Riles, Łukasz Szulc, Yerina S Ranjit, Filippo Trevisan, Cheryl Ruth R. Soriano, Jessica Vitak, and Payal Arora. 2021. "Open Science, Closed Doors? Countering Marginalization through an Agenda for Ethical, Inclusive Research in Communication." *Journal of Communication* 71 (5): 764–784.
- Francoli, Mary. 2015. *Canada Progress Report: 2012–13*. https://www.opengovpartnership.org/wp-content/uploads/2019/07/Canada_final_2012_Eng.pdf.
- Francoli, Mary. 2016. *Canada Progress Report: 2014–15*. https://www.opengovpartnership.org/wp-content/uploads/2017/05/1.Canada14-15_English_Final_o_o.pdf.
- Francoli, Mary. 2017. *Canada End of Term Report 2014–16*. https://www.opengovpartnership.org/wp-content/uploads/2017/05/Canada_EOTR_2014-2016_EN.pdf.
- Fraundorfer, M. 2017. "The Open Government Partnership: Mere Smokescreen or New Paradigm?" *Globalizations* 14 (4): 611–626. <https://doi.org/10.1080/14747731.2016.1236463>.
- Franzke, Aline, Iris Muis, and Mirko Tobias Schäfer. 2021. "Data Ethics Decision Aid (DEDA): a dialogical framework for ethical inquiry of AI and data projects in the Netherlands." *Ethics and Information Technology* (2021): 1–17.
- Franzke, Aline, Anja Bechnman, Michael Zimmer, and Charles Ess. 2019. "Internet research: Ethical guidelines 3.0." Association of Internet Researchers. <https://members.aoir.org/resources/Documents/IRE%203.0%20-%20final%20distribution%20copy%20for%20AoIR%20members.pdf>(9).
- Freeman, Christopher, and Luc Soete. 2009. "Developing Science, Technology and Innovation Indicators: What We Can Learn from the Past." *Research Policy* 38 (4): 583–589.
- Fregoso, Rosa-Linda, and Cynthia Bejarano, eds. 2010. *Terrorizing Women: Femicide in the Americas*. Durham, NC: Duke University Press.
- Fuchs, Christian. 2011. "Web 2.0, Prosumption, and Surveillance." *Surveillance & Society* 8 (3): 288–309.
- Fuller, Duncan, and Rob Kitchin, eds. 2004. *Radical Theory/critical Praxis: Making a Difference Beyond the Academy?* Vernon and Victoria: Praxis (e)Press.
- Fuller, Duncan, and Rob Kitchin, eds. 2004a. *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?* Praxis E-Press. <http://www.praxis-epress.org/rtcp/contents.html>.
- Fuller, Duncan, and Rob Kitchin. 2004b. "Radical Theory/Critical Praxis." In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 1–20. Vernon and Victoria: Praxis (e)Press.

- Fumega, Silvana. 2019a. "Open Data, Gender and Violence in Latin America." *European Public Mosaic* 7 (February): 36–47.
- Fumega, Silvana. 2019b. "Standardisation of Femicide Data Requires a Complex, Participatory Process, but Important Lessons Are Already Emerging | LSE Latin America and Caribbean". *LSE Latin America and Caribbean Blog* (blog). 8 August 2019. <https://blogs.lse.ac.uk/latamcaribbean/2019/08/08/standardisation-of-femicide-data-requires-a-complex-participatory-process-but-important-lessons-are-already-emerging/>.
- Funtowicz, Silvio O., and Jerome R. Ravetz. 1990. *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer Academic Publishers.
- Gatenby, Bev, and Maria Humphries. 2000. "Feminist Participatory Action Research: Methodological and Ethical Issues." *Women's Studies International Forum* 23 (1): 89–105. [https://doi.org/10.1016/S0277-5395\(99\)00095-3](https://doi.org/10.1016/S0277-5395(99)00095-3).
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2020. "Datasheets for Datasets." ArXiv:1803.09010 [Cs], March. <http://arxiv.org/abs/1803.09010>.
- Gerbaudo, Paolo. 2016. "From Data Analytics to Data Hermeneutics. Online Political Discussions, Digital Methods and the Continuing Relevance of Interpretive Approaches." *Digital Culture & Society* 2 (2): 95–112. <https://doi.org/10.14361/dcs-2016-0207>.
- Gewin, Virginia. 2018. "Real-Life Stories of Online Harassment -and How Scientists Got through It." *Nature* 562 (7727): 449–450. <https://doi.org/10.1038/d41586-018-07046-0>.
- Ghorashi, Halleh, and Harry Wels. 2009. "Beyond Complicity: A Plea for Engaged Ethnography." *Organizational Ethnography: Studying the Complexities of Everyday Life*, 231–253. Thousand Oaks: Sage.
- Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. 1994. *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage.
- Giddens, Anthony. 1990. *Consequences of Modernity*. Cambridge: Polity Press.
- Gilmore, Ruth Wilson. 2007. *Golden Gulag: Prisons, Surplus, Crisis, and Opposition in Globalizing California*. Berkeley, CA: University of California Press.
- Gleeson, Justin, Rob Kitchin, and Eoghan McCarthy. 2022. "Dashboards and Public Health: The Development, Impacts, and Lessons from the Irish Government COVID-19 Dashboards." *American Journal of Public Health* 112 (6): 896–893.
- Gold, Raymond. L. 1957. "Roles in Sociological Field Observations." *Social Forces* 36 (3): 217–223.
- Gonzalez Torres, Ana Paula, and Nitin Sawhney. 2023. "Role of Regulatory Sandboxes and MLOps for AI-Enabled Public Sector Services." *The Review of Socionetwork Strategies* 17: 297–318. <https://doi.org/10.1007/s12626-023-00146-y>.

- Gonzalez Torres, Ana Paula, Kaisla Kajava, and Nitin Sawhney. 2023. "Emerging AI Discourses and Policies in the EU: Implications for Evolving AI Governance." In *Artificial Intelligence Research*, edited by Anban Pillay, Edgar Jembere, and Aurore J. Gerber. SACAIR 2023. Communications in Computer and Information Science, vol. 1976. Cham: Springer. https://doi.org/10.1007/978-3-031-49002-6_1
- Góralaska, Magdalena. 2020. "Anthropology from Home: Advice on Digital Ethnography for the Pandemic Times." *Anthropology in Action* 27 (1): 46–52.
- Gordon, Sue. 2004. "Understanding Students' Experiences of Statistics in a Service Course." *Statistics Education Research Journal* 3 (1): 40–59.
- Goulet, Jean-Guy A., and Bruce Granville Miller. 2007. "Embodied Knowledge: Steps toward a Radical Anthropology of Cross-Cultural Encounters." In *Extraordinary Anthropology: Transformations in the Field*, edited by Jean-Guy A. Goulet and Bruce Granville Miller, 1–13. Lincoln, NE: University of Nebraska Press.
- Graham, Mark, and Mohammad Amir Anwar. 2019. "The Global Gig Economy: Towards a Planetary Labour Market?" *First Monday*. <https://doi.org/10.5210/fm.v24i4.9913>.
- Grant, Paul, and Lew Perren. 2002. "Small Business and Entrepreneurial Research: Meta-Theories, Paradigms and Prejudices." *International Small Business Journal* 20 (2): 185–211.
- Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. "Data Infrastructure Literacy." *Big Data & Society* 5 (2): 1–13. <https://doi.org/10.1177/2053951718786316>.
- Gray, Jonathan, Danny Lämmerhirt, and Liliana Bounegru. 2016. "Changing What Counts: How Can Citizen-Generated and Civil Society Data Be Used as an Advocacy Tool to Change Official Data Collection?" CIVICUS and Open Knowledge International. <http://papers.ssrn.com/abstract=2742871>.
- Gray, Jonathan, Liliana Bounegru, Richard Rogers, Tommaso Venturini, Donato Ricci, Axel Meunier, Sabine Niederer, Natalia Sánchez Querubín, Marc Tuters, Lucy Kimbell, and Anders Kristian Munk. 2022. "Composing Collective Inquiry with Digital Methods and Data." *Digital Culture & Education (DCE)* 14 (3): 55–86.
- Gray, Jonathan. 2016. "Datafication and Democracy: Recalibrating Digital Information Systems to Address Societal Interests." *Juncture* 23 (3): 197–201. <https://doi.org/10.1111/newe.12013>.
- Gray, Jonathan. 2018. "Three Aspects of Data Worlds." *Krisis: Journal for Contemporary Philosophy*, no. 1. <https://archive.krisis.eu/three-aspects-of-data-worlds/>.
- Gray, Mary L., and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Illustrated edition. Boston: Houghton Mifflin Harcourt.
- Grimmelikhuijsen, Stephan, and Albert Meijer. 2022. "Legitimacy of Algorithmic Decision-Making: Six Threats and the Need for a Calibrated Institutional Response." *Perspectives on Public Management and Governance* 5 (3): 232–242.

- Guston, David H. 2004. "Forget Politicizing Science. Let's Democratize Science!" *Issues in Science and Technology* 21 (1): 25–28.
- Habermas, Jürgen, and Niklas Luhmann. 1971. *Theorie der Gesellschaft oder Sozialtechnologie*. Frankfurt am Main: Suhrkamp.
- Hamilton, Malika. 2019. "Engaged Youth: Let's Talk." March 25, 2019. <https://www.opengovpartnership.org/stories/carleton-university-student-blog-series-engaged-youth-lets-talk/>.
- Hammersley, Martyn, and Paul Atkinson. 1995. *Ethnography: Principles in Practice*, second edition. New York, NY: Routledge.
- Haraway, Donna. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York, NY: Routledge.
- Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575–599.
- Harman, Grant. 2005. "Australian Social Scientists and Transition to a More Commercial University Environment." *Higher Education Research & Development* 24 (1): 79–94. <https://doi.org/10.1080/0729436052000318587>.
- Heeks, Richard, Mark Graham, Paul Mungai, Jean-Paul van Belle, and Jamie Woodcock. 2021. "Systematic Evaluation of Gig Work Against Decent Work Standards: The Development and Application of the Fairwork Framework." *The Information Society* 37 (5): 267–286. <https://doi.org/10.1080/01972243.2021.1942356>.
- Hennen, Leonhard. 2012. "Why Do We Still Need Participatory Technology Assessment?" *Poiesis & Praxis* 9: 27–41. <https://doi.org/10.1007/s10202-012-0122-5>.
- Hepp, Andreas, and Wiebke Loosen. 2019. "Molo. News: Experimentally Developing a Relational Platform for Local Journalism." *Media and Communication* 7 (4): 56–67.
- Holbrook, J. Britt. 2017. "The Future of the Impact Agenda Depends on the Reevaluation of Academic Freedom." *Palgrave Communications* 3 (39): 1–9. <https://doi.org/10.1057/s41599-017-0041-0>.
- Hunter, Savannah, Jared Joseph, Lindsay Poirier, and Nicholas Shapiro. 2021. "Carceral-Ecology OSHA Data Extraction and Cleaning Tool (version 0.0.2)." *Zenodo*. <https://doi.org/10.5281/zenodo.4825088>.
- ILDA. 2021. "Data Against Femicide: Technology in Action." YouTube, December 15, 2021, 2:51:50. <https://www.youtube.com/watch?v=BhXG7pnzNW4>.
- Iliadis, Andrew, and Federica Russo. 2016. "Critical Data Studies: An Introduction." *Big Data & Society* 3 (2): 2053951716674238.
- Immigration, Refugees and Citizenship Canada (IRCC). 2017. *Evaluation of the Settlement Program*. Research and Evaluation Branch, IRCC. <https://www.canada.ca/content/dam/ircc/documents/pdf/english/evaluation/e2-2016-settlement-en.pdf>.
- Jaeger, Paul T., and John Carlo Bertot. 2010. "Transparency and Technological Change: Ensuring Equal and Sustained Public Access to Government Information." *Government Information Quarterly* 27 (4): 371–376. <https://doi.org/10.1016/j.giq.2010.05.003>.

- Jahn, Thomas, Matthias Bergmann, and Florian Keil. 2012. "Transdisciplinarity: Between Mainstreaming and Marginalization." *Ecological Economics* 79 (July): 1–10. <https://doi.org/10.1016/j.ecolecon.2012.04.017>.
- Jahoda, Marie, Paul F. Lazarsfeld, and Hans Zeisel. 1974. *Marienthal: The Sociology of an Unemployed Community*. New Brunswick, NJ: Transaction Publishers.
- Jain, Sumeet, and Sushrut Jadhav. 2009. "Pills That Swallow Policy: Clinical Ethnography of a Community Mental Health Program in Northern India." *Transcultural Psychiatry* 46 (1): 60–85.
- Jansen, Julia, Sue-Yen Tjong Tjin Tai, Jos van den Broek, and Jasper Deuten. 2020. "Stad Zoekt Toga – Universiteiten en Hogescholen als Structurele Kennispartner voor Gemeenten." The Hague: Rathenau Institute.
- Janský, Petr. 2020. "European Banks and Tax Havens: Evidence from Country-by-Country Reporting." *Applied Economics* 52 (54): 5967–5985. <https://doi.org/10.1080/00036846.2020.1781773>.
- Jaton, Florian. 2021. *The Constitution of Algorithms: Ground-Truthing, Programming, Formulating*. Cambridge, MA: MIT Press.
- Jensen, Torben Elgaard, Andreas Birkbak, Anders Koed Madsen, and Anders Kristian Munk. 2021. "Participatory Data Design: Acting in a Digital World." In *Making & Doing: Activating STS through Knowledge Expression and Travel*, edited by Gary Lee Downey and Teun Zuiderent-Jerak, 117–136. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/11310.001.0001>.
- Jensenius, Francesca, Mala Htun, David J. Samuals, David A. Singer, Adria Lawrence, and Michael Chwe. 2018. "The Benefits and Pitfalls of Google Scholar." *American Political Science Association* 51 (4): 820–824. <https://doi.org/10.1017/S104909651800094X>.
- Kaba, Mariame. 2021. *We Do This 'til We Free Us: Abolitionist Organizing and Transforming Justice*. Abolitionist Papers Series. Chicago, IL: Haymarket Books.
- Kahne, Joseph, Ellen Middaugh and Danielle Allen. 2014. *Youth, New Media, and the Rise of Participatory Politics*. March 19, 2014. Oakland, CA: Youth and Participatory Politics Research Network. https://clalliance.org/wp-content/uploads/files/ypp_workinpapers_paper01_1.pdf.
- Kania, John, Junious Williams, Paul Schmitz, Sheri Brady, Mark Kramer, and Jennifer Splansky Juster. 2022. "Centering Equity in Collective Impact." *Stanford Social Innovation Review* 20 (1): 38–45.
- Keen, Andrew. 2007. *The Cult of the Amateur: How Blogs, MySpace, YouTube, and the Rest of Today's User-Generated Media Are Destroying Our Economy, Our Culture, and Our Values*. New York, NY: Doubleday.
- Kennedy, Heather and Giles Moss. 2015. "Known or Knowing Publics? Social Media Data Mining and the Question of Public Agency." *Big Data & Society* 2 (2). <https://doi.org/10.1177/2053951715611145>.

- Kim, Gang-Hoon, Silvana Trimi, and Ji-Hyong Chung. 2014. "Big-Data Applications in the Government Sector." *Communications of the ACM* 57 (3): 78–85. <https://doi.org/10.1145/2500873>.
- Kim, Nadia Y. 2021. *Refusing Death: Immigrant Women and the Fight for Environmental Justice in LA*. Stanford, CA: Stanford University Press.
- Kitchin, Rob, and Tracey P. Lauriault. 2014. "Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work." SSRN Scholarly Paper ID 2474112. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=2474112>.
- Kitchin, Rob, Denis Linehan, Cian O'Callaghan, and Philip Lawton. 2013. "Public Geographies and Social Media." *Dialogues in Human Geography* 3 (1): 56–72.
- Kitchin, Rob, Sophia Maalsen, and Gavin McArdle. 2016. "The Praxis and Politics of Building Urban Dashboards." *Geoforum* 77: 93–101.
- Kitchin, Rob, Tracey P. Lauriault, and Gavin McArdle. 2017. "Data and the City." In *Data and the City*, edited by Rob Kitchin, Tracey P. Lauriault, and Gavin McArdle, 1–13. Abingdon: Routledge.
- Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Los Angeles, CA: Sage.
- Kitchin, Rob. 2014a. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1 (1): 1–12.
- Kitchin, Rob. 2015. "Positivistic Geography." In *Approaches to Human Geography*, 2nd edition, edited by Stuart Aitken, and Gill Valentine, 23–34. London: Sage.
- Kitchin, Rob. 2021. *Data Lives: How Data Are Made and Shape Our Lives*. Bristol: Bristol University Press.
- Kitchin, Rob. 2022. *The Data Revolution: A Critical Approach to Big Data, Open Data, and Data Infrastructures*, 2nd edition. London: Sage.
- Klein, Julie. 2004. "Prospects for Transdisciplinarity." *Futures* 36 (4): 515–26. <https://doi.org/10.1016/j.futures.2003.10.007>.
- Klievink, Bram, Bart-Jan Romijn, Scott Cunningham, and Hans de Bruijn. 2017. "Big Data in the Public Sector: Uncertainties and Readiness." *Information Systems Frontiers* 19 (2): 267–83. <https://doi.org/10.1007/s10796-016-9686-2>.
- Knorr-Cetina, Karin D. 1981. "The Micro-Sociological Challenge of Macro-Sociology: Towards a Reconstruction of Social Theory and Methodology." In *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro-Sociologies*, edited by Aaron Victor Cicourel, and Karin D. Knorr-Cetina, 1–47. Boston: Routledge & Kegan Paul.
- Kohonen, Matti, and Francine Mestrum, eds. 2008. *Tax Justice: Putting Global Inequality on the Agenda*. London: Pluto Press.
- Kolkman, Daan. 2020. "The Usefulness of Algorithmic Models in Policy Making." *Government Information Quarterly* 37 (3): 101488.

- Kolkman, Daan. 2022. "The (In) Credibility of Algorithmic Models to Non-Experts." *Information, Communication & Society* 25 (1): 93–109.
- Kolkman, Daan. A. 2016. "Models in Policy Making." PhD Thesis, University of Surrey.
- Kolkman, Daan. A., Paolo Campo, Tina Balke-Visser, and Nigel Gilbert. 2016. "How to Build Models for Government: Criteria Driving Model Acceptance in Policymaking." *Policy Sciences* 49 (4): 489–504.
- König, René. 2012. "Wikipedia. Between Lay Participation and Elite Knowledge Representation." *Information, Communication & Society*, October. <https://doi.org/10.1080/1369118X.2012.734319>.
- Kuhn, Thomas. S. 1963. "The Function of Dogma in Scientific Research." In *Scientific Change: Historical Studies in the Intellectual, Social, and Technical Conditions for Scientific Discovery and Technical Invention, From Antiquity to the Present*, edited by Alistair Cameron Crombie, 347–369. New York, NY: Basic Books.
- Labaree, Robert V. 2002. "The Risk of 'Going Observationalist': Negotiating the Hidden Dilemmas of Being an Insider Participant Observer." *Qualitative Research* 2 (1): 97–122.
- Lagarde y de los Ríos, Marcela. 2010. "Preface: Feminist Keys for Understanding Femicide: Theoretical, Political and Legal Construction." In *Terrorizing Women: Femicide in the Americas*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, xi–xxv. Durham, NC: Duke University Press.
- Laidlaw, Sam. 2012. *Report of the Laidlaw Inquiry: Inquiry into the Lessons Learned for the Department for Transport from the Intercity West Coast Competition*. London: The Stationery Office.
- Lampland, Martha, and Star, Susan Leigh. 2009. *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.
- Landström, Hans, and Franz Lohrke. 2010. *Historical Foundations of Entrepreneurship Research*. Cheltenham: Edward Elgar Publishing.
- Landström, Hans, and Mats Benner. 2010. "Entrepreneurship Research: A History of Scholarly Migration." In *Historical Foundations of Entrepreneurship Research*, edited by Hans Landström and Franz Lohrke, 15–45. Cheltenham: Edward Elgar Publishing.
- Latour, Bruno, and Steve Woolgar. 1979. *Laboratory Life*. Princeton, NJ: Princeton University Press.
- Latour, Bruno, and Steve Woolgar. 2013. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.
- Latour, Bruno. 1988. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.
- Latour, Bruno. 1999. "Circulating Reference: Sampling the Soil in the Amazon Forest." In *Pandora's Hope: Essays on the Reality of Science Studies*, 24–79. Cambridge, MA: Harvard University Press.

- Lazarsfeld, Paul, F. 1975. "The Policy Science Movement (An Outsider's View)." *Policy Sciences* 6 (3): 211–222.
- Leaman, Jeremy, and Attiya Waris, eds. 2013. *Tax Justice and the Political Economy of Global Capitalism, 1945 to the Present*. New York, NY: Berghahn Books.
- Lee, Yong S. 1996. "'Technology Transfer' and the Research University: A Search for the Boundaries of University-Industry Collaboration." *Research Policy* 25 (6): 843–863. [https://doi.org/10.1016/0048-7333\(95\)00857-8](https://doi.org/10.1016/0048-7333(95)00857-8).
- Lee, Yong S. 2000. "The Sustainability of University-Industry Research Collaboration: An Empirical Assessment." *The Journal of Technology Transfer* 25 (2): 111–133.
- Lehtiniemi, Tuukka. 2023. "Contextual Social Valences for Artificial Intelligence: Anticipation that Matters in Social Work." *Information, Communication & Society*: 1–16.
- Leonelli, Sabina, Brian Rappert, and Gail Davies. 2017. "Data Shadows: Knowledge, Openness, and Absence." *Science, Technology & Human Values* 42 (2): 191–202.
- Lewis, Tamika, Seeta Peña Gangadharan, Mariella Saba, and Tawana Petty. 2018. *Digital Defense Playbook: Community Power Tools for Reclaiming Data*. Our Data Bodies. http://www.odbproject.org/wp-content/uploads/2019/03/ODB_DDP_HighRes_Single.pdf.
- Lightman, Naomi, and Luann Good Gingrich. 2018. "Measuring Economic Exclusion for Racialized Minorities, Immigrants and Women in Canada: Results from 2000 and 2010." *Journal of Poverty* 22 (5): 398–420. <https://doi.org/10.1080/10875549.2018.1460736>.
- Lima, José Júlio Ferreira, Ricardo Bruno Santos, Lucinda Freitas de Assis Sena, Camila Leandra Araújo. 2015. "Estrutura Social e Organização Social da Região Metropolitana de Belém." In *Belém: Transformações na ordem Urbana*, edited by Ana Cláudia Duarte Cardoso, and José Júlio Ferreira Lima, 145–172. Rio de Janeiro: Observatório das metrópoles: Letra. capital. <http://livroaberto.ufpa.br/jspui/handle/prefix/372>.
- Livingstone, Sonia. 2019. "Audiences in an Age of Datafication: Critical Questions for Media Research." *Television & New Media* 20 (2): 170–183.
- Loukissas, Yanni Alexander. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Lovink, Geert, and Nathaniel Tkacz, eds. 2011. *Critical Point of View: A Wikipedia Reader*. Amsterdam: Institute of Network Cultures. <http://networkcultures.org/wpmu/portal/publications/inc-readers/critical-point-of-view-a-wikipedia-reader/>.
- Luján Pinelo, Aleida. 2018. "A Theoretical Approach to the Concept of Femi(n)icide." *Philosophical Journal of Conflict and Violence* 2 (1): 41–63. <https://doi.org/10.22618/TP.PJCV.20182.1.171003>.
- Lupton, Deborah, and Mike Michael. 2017. "'Depends on Who's Got the Data': Public Understandings of Personal Digital Dataveillance." *Surveillance and Society* 15 (2): 254–268.

- MacGregor, Jennifer CD., Casey L. Oliver, Barbara J. MacQuarrie and C. Nadine Wathen. 2019. "Intimate Partner Violence and Work: A Scoping Review of Published Research." *Trauma, Violence, & Abuse*, 22 (4): 717–727. <https://doi.org/10.1177/1524838019881746>.
- Maciejewski, Mariusz. 2017. "To Do More, Better, Faster and More Cheaply: Using Big Data in Public Administration." *International Review of Administrative Sciences* 83 (1_suppl): 120–35. <https://doi.org/10.1177/0020852316640058>.
- Madrigal, Sonia, Ivonne Ramírez Ramírez, María Salguero, and Helena Suárez Val. 2019. "Monitoring, Recording, and Mapping Femicide – Experiences from Mexico and Uruguay." In *Femicide Volume XII: Living Victims of Femicide*, edited by Helen Hemblade and Helena Gabriel, 67–73. Vienna: ACUNS. http://femicide-watch.org/sites/default/files/Femicide%20XII_o.pdf#page=73.
- Mager, Astrid. 2012. "Algorithmic Ideology. How Capitalist Society Shapes Search Engines." *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2012.676056>
- Maier, Daniel, et al. 2018. "Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology." *Communication Methods and Measures*, 12 (2–3): 93–118. <https://doi.org/10.1080/19312458.2018.1430754>.
- Manovich, Lev. 2011. "Trending: The Promises and the Challenges of Big Social Data." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 460–475. Minneapolis, MI: University of Minnesota Press.
- Markham, Annette N. 2006. "Ethic as Method." *Journal of Information Ethics*, 15(2): 37–54. DOI:10.3172/JIE.15.2.37.
- Markham, Annette N. 2017. "Troubling the Concept of Data in Digital Qualitative Research." In *Handbook of Qualitative Data Collection*, edited by Uwe Flick, 511–523. London: Sage.
- Markham, Annette, and Elizabeth Buchanan. 2012. "Ethical Decision-Making and Internet Research: Recommendations from the AOIR Ethics Working Committee (Version 2.0)." Association of Internet Research. <http://aoir.org/reports/ethics2.pdf>.
- Markham, Annette. 2019. "Critical Pedagogy as a Response to Datafication." *Qualitative Inquiry* 25 (8): 754–760.
- Markham, Annette. 2021. "The Limits of the Imaginary: Challenges to Intervening in Future Speculations of Memory, Data, and Algorithms." *New Media and Sociology* 23 (2): 382–405.
- Martin, Aryn, and Michael Lynch. 2009. "Counting Things and People: The Practices and Politics of Counting." *Social Problems* 56 (2): 243–266. <https://doi.org/10.1525/sp.2009.56.2.243>.
- Martín-Barbero, Jesus, and Omar Rincón. 2019. "Mapa insomne 2017: Ensayos sobre el sensorium contemporáneo. Un mapa para investigar la mutación cultural."

- In *Un nuevo mapa para investigar la mutación cultural: Diálogo con la propuesta de Jesús Martín-Barbero*. Quito: CIESPAL.
- Martín-Barbero, Jesus. 2011. "La pertenencia en el horizonte de las nuevas tecnologías y de la sociedad de la comunicación." In *Sentido de pertenencia en sociedades fragmentadas: América Latina desde una perspectiva global* by Ana Sojo, and Martín Hopenhayn, 105–127. Buenos Aires: Siglo Veintiuno.
- Masso, Anu, Katrin Tiidenberg, and Andra Siibak. 2020. "Kuidas mõista andmestunud maailma? Metodoloogiline teejuht." TLÜ Kirjastus. <https://www.tlu.ee/pood/home/325-kuidas-moista-andmestunud-maailma-metodoloogiline-teejuht.html>.
- Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York, NY: Houghton Mifflin Harcourt.
- McCosker, Anthony, Jane Farmer and Arezou Soltani Panah. 2020. *Community Responses to Family Violence: Charting Policy Outcomes using Novel Data Sources, Text Mining and Topic Modelling*. Melbourne: Swinburne University of Technology. <https://apo.org.au/node/278041>.
- McCosker, Anthony, Xiaofang Yao, Kath Albury, Alexia Maddox, Jane Farmer and Julia Stoyanovich. 2022. "Developing Data Capability with Non-Profit Organisations Using Participatory Methods." *Big Data & Society* 9 (1). <https://doi.org/10.1177/20539517221099882>.
- McIntyre, Lee. 2018. *Post-Truth*. Cambridge, MA: MIT Press.
- McNealy, Jasmine. 2020. "An Ecological Approach to Data Governance." Data & Society Research Institute. https://youtu.be/jB5_NrdWH7k.
- McNiff, Jean. 2016. *You and Your Action Research Project*. London: Routledge.
- Mehta, Chinar. 2021. "What Does the Abandoned Public Toilet Tell Us?" *FemLab. Co* (blog). 2021. <https://femlab.co/2021/05/21/what-does-the-abandoned-public-toilet-tell-us/>.
- Meier, Patrick. 2015. *Digital Humanitarians: How Big Data Is Changing the Face of Humanitarian Response*, 1st ed. Oxfordshire: Routledge. <https://doi.org/10.1201/b18023>.
- Meijer, Albert, Alex Ingrams, and Stavros Zouridis. 2022. *Public Management in an Information Age: Towards Strategic Public Information Management*. Bloomsbury Publishing.
- Meijer, Albert, and Erna Ruijter. 2021. *Code Goed Digitaal Openbaar Bestuur*. Government of the Netherlands. [https://www.rijksoverheid.nl/documenten/rapporten/2021/04/30/code-goed-digitaal-openbaar-bestuur](https://www.rijksoverheid.nl/documenten/rappporten/2021/04/30/code-goed-digitaal-openbaar-bestuur).
- Meijer, Albert, Mirko Tobias Schäfer, and Martiene Branderhorst. 2019. "Principes Voor Goed Lokaal Bestuur in de Digitale Samenleving: Een Aanzet tot een Normatief Kader." *Bestuurswetenschappen* 73 (4): 8–23.

- Meijer, Albert. 2018. "Datapolis: A Public Governance Perspective on 'Smart Cities.'" *Perspectives on Public Management and Governance* 1 (3): 195–206.
- Mejias, Ulises A., and Nick Couldry. 2019. "Datafication." *Internet Policy Review* 8 (4). <https://doi.org/10.14763/2019.4.1428>.
- Melchior, Jillian Kay. 2018. "Opinion | Fake News Comes to Academia." *Wall Street Journal*, October 5, 2018, sec. Opinion. <https://www.wsj.com/articles/fake-news-comes-to-academia-1538520950>.
- Melody, William, H. and Mansell, Robin, E. 1983. "The Debate over Critical vs. Administrative Research: Circularity or Challenge." *Journal of Communication*, 33(3): 103–116.
- Merry, Sally Engle. 2016. *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*. Chicago, IL: University of Chicago Press.
- Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago Press.
- Micheli, Marina, Marisa Ponti, Max Craglia, and Anna Berti Suman. 2020. "Emerging Models of Data Governance in the Age of Datafication." *Big Data & Society* 7 (2): 205395172094808. <https://doi.org/10.1177/2053951720948087>.
- Miedema, Frank. 2022. *Open Science: The Very Idea*. Dordrecht: Springer.
- Milan, Stefania, and Lonneke van der Velden. 2016. "The Alternative Epistemologies of Data Activism." *Digital Culture & Society* 2 (2): 57–74.
- Ministerie van Economische Zaken, and Ministerie van Onderwijs, Cultuur en Wetenschap. 2013. *Nederlandse Oplossingen Voor Wereldwijde Problemen*. Rotterdam: Vijfkeerblauw. <https://www.rvo.nl/sites/default/files/2013/11/Nederlandse-oplossingen-voor-wereldwijde-uitdagingen-Topsectoren-en-Horizon-2020.pdf>.
- Ministerie van Infrastructuur en Milieu. 2016. *Omgevingswet in het Kort: Ruimte voor Ontwikkeling, Waarborgen voor Kwaliteit*. Interdepartementale Programmadirectie Eenvoudig Beter. www.omgevingswetportaal.nl.
- Mitchell, Don. 2004. "Radical Scholarship: A Polemic on Making a Difference Outside the Academy." In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 21–30. Vernon and Victoria: Praxis (e)Press.
- Mobjörk, Malin. 2010. "Consulting versus Participatory Transdisciplinarity: A Refined Classification of Transdisciplinary Research." *Futures* 42 (8): 866–73. <https://doi.org/10.1016/j.futures.2010.03.003>.
- Montoya-Barthelemy, Andre. 2019. "The Occupational Health of Prison Inmates: An Ignored Population and an Opportunity." *Journal of Occupational and Environmental Medicine* 61 (2): e74–76. <https://doi.org/10.1097/JOM.0000000000001514>.
- Mosley, Mark and Michael Brackett. 2010. *The DAMA Guide to the Data Management Body of Knowledge (DAMA-DMBOK Guide)*. 1st edition. Bradley Beach, NJ: Technics Publications.

- Mounk, Yascha. 2018. "What an Audacious Hoax Reveals About Academia." *The Atlantic*. October 5, 2018. <https://www.theatlantic.com/ideas/archive/2018/10/new-sokal-hoax/572212/>.
- Mourat, Robin de, Donato Ricci, and Bruno Latour. 2020. "How Does a Format Make a Public?" In *Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access*, edited by Martin Paul Eve and Jonathan Gray, 103–112. Cambridge, MA: MIT Press.
- Mouton, Johann, and Hedrik Christoffel Marais. 1988. *Basic Concepts in the Methodology of the Social Sciences*. Pretoria: HSRC Publishers.
- Munk, Anders Kristian, Axel Meunier, and Tommaso Venturini. 2019. "Data Sprints: A Collaborative Format in Digital Controversy Mapping." In *DigitalSTS: A Field Guide for Science and Technology Studies*, edited by Janet Vertesi and David Ribes, 472–496. Princeton, NJ: Princeton University Press.
- Murthy, Dhiraj. 2008. "Digital Ethnography: An Examination of the Use of New Technologies for Social Research." *Sociology* 42 (5): 837–855.
- Nader, Laura. 1972. "Up the Anthropologist: Perspectives Gained from Studying Up." In *Reinventing Anthropology*, edited by Dell Hymes, 284–311. New York, NY: Vintage.
- Nelson, Andrew J. 2005. "Cacophony or Harmony? Multivocal Logics and Technology Licensing by the Stanford University Department of Music." *Industrial and Corporate Change* 14 (1): 93–118.
- Nentwich, Michael, and René König. 2012. *Cyberscience 2.0. Research in the Age of Digital Social Networks*. Vol. 11. Interactiva. Frankfurt: Campus.
- Newquist, Harvey. P. 1994. *The Brain Makers*. Indianapolis, IN: Sams Publishing.
- Nguyen, Dennis and Karin van Es. 2024 [Forthcoming]. "Exploring the Value of Computational Methods for Journalistic Self-Reflection and 'Accountability.'" *Journalism Studies*.
- Nieuwenhuizen, Wouter. 2020. *How Working Digitally During the Covid-19 Pandemic Does Not Help to Transform the Public Sector: A Digital Organisational Ethnography*. Master's thesis, Utrecht University.
- Nogrady, Bianca. 2021. "'I Hope You Die': How the COVID Pandemic Unleashed Attacks on Scientists." *Nature* 598 (7880): 250–253. <https://doi.org/10.1038/d41586-021-02741-x>.
- Nowotny, Helga. 2003. "Democratising Expertise and Socially Robust Knowledge." *Science and Public Policy* 30 (3): 151–156.
- O'Reilly, Tim. 2005. "What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software." September 30, 2005. <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Broadway Books.

- Oostrom, Frits. 2007. "Markt en ziel. De universiteit als instituut van de verbeelding." *Ons Erfdeel* 50 (4): 54–65.
- OSHA. n.d. "State Plans | Occupational Safety and Health Administration." Accessed November 5, 2021. <https://www.osha.gov/stateplans/>.
- Ottawa Local Immigration Partnership (OLIP). 2016. "Better Together: Cultivating Equitable and Inclusive Organizations in Ottawa: Summary Report on Phase I of Ottawa's Equity Project." <http://olip-plio.ca/wp-content/uploads/2013/07/OLIP-Equity-Report-EN-Low.pdf>.
- Pangrazio, Luci and Neil Selwyn. 2019. "Personal Data Literacies': A Critical Literacies Approach to Enhancing Understandings of Personal Digital Data." *New Media & Society* 21 (2): 419–437. <https://doi.org/10.1177/1461444818799523>.
- Paré, Daniel. 2018. *Ontario Progress Report 2017*. https://www.opengovpartnership.org/wp-content/uploads/2018/10/Ontario_Final-Report_2017.pdf.
- Paré, Daniel. 2021. *Canada Design Report 2018–2020*. https://www.opengovpartnership.org/wp-content/uploads/2021/07/Canada_Design_Report_2018-2020_EN.pdf
- Pastran, Camila. 2016. *Open Government: A Solid Step Towards Peace Agreements*. <https://www.opengovpartnership.org/stories/open-government-a-solid-step-towards-peace-agreements/>.
- Patil, Dhanurjay J. 2011. *Building Data Science Teams*. O'Reilly Media, Inc.
- Pattnaik, Pinaki Nandan, and Satyendra C. Pandey. 2014. "University Spinoffs: What, Why, and How?" *Technology Innovation Management Review* 4 (12): 44–50.
- Pegado, Rosielle and Claudio Blanco, Jackson Roehrig, Carla Carocha, and Francisco Costa. 2014. *Risco de Cheia e Vulnerabilidade: Uma Abordagem às Inundações Urbanas de Belém/Pará/Brasil*. *Revista Territorium*. 21. 71. https://doi.org/10.14195/1647-7723_21_6.
- Penteado, Antonio Rocha. 1968. *Belém: (estudo de geografia urbana)*. Belém: UFPA. 2 v. (Coleção amazônica. Série José Veríssimo). Available at <http://livroaberto.ufpa.br/jspui/handle/prefix/43>.
- Perren, Lew, and Monder Ram. 2004. "Case-Study Method in Small Business and Entrepreneurial Research: Mapping Boundaries and Perspectives." *International Small Business Journal* 22 (1): 83–101.
- Phillips, Jeanet and Penny Vandenbroek. 2014. *Domestic, Family and Sexual Violence in Australia: An Overview of the Issues*. Canberra: Parliamentary Library, Department of Parliamentary Services. http://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1415/ViolenceAust.
- Pielke Jr, Roger A. 2007. *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge: Cambridge University Press.
- Pielke, Roger. A. 1999. "Who Decides? Forecasts and Responsibilities in the 1997 Red River Flood." *Applied Behavioral Science Review* 7 (2): 83–101.

- Pierre, Jennifer, Roderic Crooks, Morgan Currie, Britt Paris, and Irene Pasquetto. 2021. "Getting Ourselves Together: Data-Centered Participatory Design Research & Epistemic Burden." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–11. CHI '21. New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445103>.
- Pilegaard, Morten, Peter W. Moroz, and Helle Neergaard. 2010. "An Auto-Ethnographic Perspective on Academic Entrepreneurship: Implications for Research in the Social Sciences and Humanities." *The Academy of Management Perspectives* 24 (1): 46–61.
- Plantin, Jean-Christophe. 2019. "Data Cleaners for Pristine Datasets: Visibility and Invisibility of Data Processors in Social Science." *Science, Technology, & Human Values* 44 (1): 52–73. <https://doi.org/10.1177/0162243918781268>.
- Pluckrose, Helen, James A. Lindsay, and Peter Boghossian. 2018. "Academic Grievance Studies and the Corruption of Scholarship." *Areo*. October 2, 2018. <https://areomagazine.com/2018/10/02/academic-grievance-studies-and-the-corruption-of-scholarship/>.
- Poderi, Giacomo, Hanna Hasselqvist, Andrea Capaccioli, Cristian Bogdan, and Vincenzo D'Andrea. 2020. "Matters of Concerns and User Stories: Ontological and Methodological Considerations for Collaborative Design Processes." *CoDesign* 16 (3): 220–232. <https://doi.org/10.1080/15710882.2018.1557694>.
- Poirier, Lindsay. 2021. "Reading Datasets: Strategies for Interpreting the Politics of Data Signification." *Big Data & Society* 8 (2): 20539517211029320. <https://doi.org/10.1177/20539517211029322>.
- Polanyi, Michael. 1966. *The Tacit Dimension*. Chicago, IL: University of Chicago Press.
- Powell, Walter W., Kenneth W. Koput, and Laurel Smith-Doerr. 1996. "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology." *Administrative Science Quarterly* 41 (1): 116–45. <https://doi.org/10.2307/2393988>.
- Press, Gil. 2013. *A Very Short History of Data Science*. Forbes. <https://forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/>.
- Puig de la Bellacasa, María. 2017. *Matters of Care: Speculative Ethics in More than Human Worlds*. Minneapolis, MN: University of Minnesota Press. <https://www.jstor.org/stable/10.5749/j.ctt1mmfspt>.
- Puschmann, Cornelius, and Jean Burgess. 2014. "Metaphors of Big Data." *International Journal of Communication* 8 (0): 20.
- Radin, Joanna. 2017. "Digital Natives': How Medical and Indigenous Histories Matter for Big Data." *Osiris* 32 (1): 43–64. <https://doi.org/10.1086/693853>.
- Rani, Uma, and Marianne Furrer. 2020. "Algorithmic Management of Work and Workers." *Competition & Change* 25 (2): 212–236. <https://doi.org/10.1177/1024529420905187>.

- Redden, Joanna. 2018. "Democratic Governance in an Age of Datafication: Lessons from Mapping Government Discourses and Practices." *Big Data & Society* 5 (2): 205395171880914. <https://doi.org/10.1177/2053951718809145>.
- Reilly, Katherine. 2020. "The Challenge of Decolonizing Big Data through Citizen Data Audits." *Dataactive*. <https://data-activism.net/2020/04/bigdatasur-the-challenge-of-decolonizing-big-data-through-citizen-data-audits-1-3/>.
- Reiter, Dan. 2016. "Citation Count Data and Faculty Promotion." *Duck of Minerva*. February 18, 2016. <https://www.duckofminerva.com/2016/02/citation-count-data-and-faculty-promotion.html>.
- Rettberg, Jill Walker. 2020. "Situated Data Analysis: A New Method for Analyzing Encoded Power Relationships in Social Media Platforms and Apps." *Humanities and Social Sciences Communications* 7 (1): 1–13. <https://doi.org/10.1057/s41599-020-0495-3>.
- Rheingold, Howard. 2003. *Smart Mobs: The Next Social Revolution*. Revised edition. Cambridge, MA: Basic.
- Ribes, David, and Steven J Jackson. 2013. "Data Bite Man: The Work of Sustaining a Long-Term Study." In *"Raw Data" Is an Oxymoron*, edited by Lisa Gitelman, 147–66. Cambridge, MA: MIT Press.
- Richterich, Annika. 2018. *The Big Data Agenda: Data Ethics and Critical Data Studies*. London: University of Westminster Press.
- Rincón, Omar. 2018. "Mutaciones bastardas de la comunicación." *MATRIZES* 12 (1): 65–78.
- Rincón, Omar. 2019. "Mi invención: Sobre el mapa para comprender el sensorium de la contemporaneidad." In *Un nuevo mapa para investigar la mutación cultural: Diálogo con la propuesta de Jesús Martín-Barbero*. Quito: CIESPAL.
- Roberts, Steven, Helene Snee, Christine Hine, Yvette Morey, and Hayley Watson, eds. 2016. *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*. London: Springer.
- Rodrigues, Edmilson Brito. 1996. *Aventura Urbana: Urbanização, Trabalho e Meio Ambiente em Belém*. Belém: Editora.
- Rogers, Richard. 2013. *Digital Methods*. Cambridge, MA: MIT Press.
- Rose, Gillian. 1997. "Situating Knowledges: Positionality, Reflexivity and Other Tactics." *Progress in Human Geography* 21, 305–320.
- Rosenblat, Alex, and Luke Stark. 2015. "Uber's Drivers: Information Asymmetries and Control in Dynamic Work." *International Journal of Communication* 10 (27): 3758–3784. <https://doi.org/10.2139/ssrn.2686227>.
- Rottenburg, Richard, Sally E. Merry, Sung-Joon Park, and Johanna Mugler, eds. 2015. *The World of Indicators: The Making of Governmental Knowledge through Quantification*. Cambridge: Cambridge University Press.

- Ruckenstein, Minna, Sarah Pink, Melisa Duque, Blanca Callén. 2024. "Anticipatory Repair: Reframing Breakage through a Futures Agenda." Interview with Minna Ruckenstein and Sarah Pink. In *Disena* 24, January 24. <https://doi.org/10.7764/disena.24>.
- Ruijter, Erna, Justien Dingelstad, and Albert Meijer. 2023. "Studying Complex Systems Through Design Interventions Probing Open Government Data Ecosystems in the Netherlands." *Public Management Review* 25 (1): 129–149.
- Ruijter, Erna. 2021. "Designing and Implementing Data Collaboratives: A Governance Perspective." *Government Information Quarterly* 38 (4): 101612.
- Salesses, Matthew. 2021. *Craft in the Real World: Rethinking Fiction Writing and Workshopping*. New York, NY: Catapult.
- Sandoval, Cueponcaxochitl D. Moreno, Rosalva Mojica Lagunas, Lydia T Montelongo, and Marisol Juárez Díaz. 2016. "Ancestral Knowledge Systems: A Conceptual Framework for Decolonizing Research in Social Science." *AlterNative: An International Journal of Indigenous Peoples* 12 (1): 18–31.
- Santos, Boaventura de Souza. 2015. *Epistemologies of the South: Justice against epistemicide*. London: Routledge.
- Saunders, Tom and Geoff Mulgan. 2017. *Governing with Collective Intelligence*. London: Nesta. <https://www.nesta.org.uk/report/governing-with-collective-intelligence/>.
- Schäfer, Mirko Tobias, and Karin van Es. 2017. "New Brave World." In *The Datafied Society: Studying Culture Through Data*, edited by Mirko Tobias Schäfer and Karin van Es, 13–22. Amsterdam: Amsterdam University Press.
- Schäfer, Mirko Tobias, Karin van Es, and Iris Muis. 2023. "Investigating the Datafied Society: Entrepreneurial Research as an Approach." In *Situating Data: Inquiries in Algorithmic Culture*, edited by Karin van Es and Nanna Verhoeff, 267–272. Amsterdam: Amsterdam University Press.
- Schäfer, Mirko Tobias, Karin van Es, and Iris Muis. 2023. "Investigating the Datafied Society. Entrepreneurial Research as Approach." In *Situating Data: Inquiries in Algorithmic Culture*, edited by Karin van Es and Nanna Verhoeff, 267–272. Amsterdam: Amsterdam University Press.
- Schmidt, Laura, Thomas Falk, Marianna Siegmund-Schultze, and Joachim H. Spangenberg. 2020. "The Objectives of Stakeholder Involvement in Transdisciplinary Research. A Conceptual Framework for a Reflective and Reflexive Practise." *Ecological Economics* 176 (October): 106751. <https://doi.org/10.1016/j.ecolecon.2020.106751>.
- Schnapp, Kai-Uwe. 2017. "Das Hamburger 'Projektbüro Angewandte Sozialforschung.'" *Zeitschrift für Politikwissenschaft* 27 (4): 529–547.
- Seabrooke, Leonard, and Duncan Wigan. 2015a. "How Activists Use Benchmarks: Reformist and Revolutionary Benchmarks for Global Economic Justice." *Review of International Studies* 41 (5): 887–904. <https://doi.org/10.1017/S0260210515000376>.

- Seabrooke, Leonard, and Duncan Wigan. 2015b. "Powering Ideas through Expertise: Professionals in Global Tax Battles." *Journal of European Public Policy* 23 (3): 357–374. <https://doi.org/10.1080/13501763.2015.1115536>.
- Seaver, Nick. 2017. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society* 4 (2): 1–12.
- Seaver, Nick. 2021. "Care and Scale: Decorrelative Ethics in Algorithmic Recommendation." *Cultural Anthropology* 36 (3): 509–537. <https://doi.org/10.14506/ca36.3.11>.
- Sefton-Green, Julian and Luci Pangrazio. 2021. "The Death of the Educative Subject? The Limits of Criticality Under Datafication." *Educational Philosophy and Theory*. <https://doi.org/10.1080/00131857.2021.1978072>.
- Segato, Rita Laura. 2010. "Territory, Sovereignty, and Crimes of the Second State: The Writing on the Body of Murdered Women." In *Terrorizing Women*, edited by Rosa-Linda Fregoso and Cynthia Bejarano, translated by Sara Koopman, 70–92. Durham, NC: Duke University Press. <https://doi.org/10.1215/9780822392644-005>.
- Selwyn, Neil, and Luci Pangrazio. 2018. "Doing Data Differently? Developing Personal Data Tactics and Strategies amongst Young Mobile Media Users." *Big Data & Society* 5 (1): 2053951718765021. <https://doi.org/10.1177/2053951718765021>.
- Shane, Scott Andrew. 2004. *Academic Entrepreneurship: University Spinoffs and Wealth Creation*. Cheltenham: Edward Elgar Publishing.
- Shapin, Steven, and Simon Schaffer. 1985. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.
- Shirky, Clay. 2008. *Here Comes Everybody: The Power of Organizing Without Organizations*. London: Penguin.
- Siffels, Lotje, David van den Berg, Mirko Tobias Schäfer, and Iris Muis. 2022. "Public Values and Technological Change: Mapping How Municipalities Grapple with Data Ethics." In *New Perspectives in Critical Data Studies: The Ambivalences of Data Power*, edited by Andreas Hepp, Juliane Jarke, and Leif Kramp, 243–265. Cham: Palgrave Macmillan.
- Silva, Rogério Renato, Max Gasparini, Elis Alquezar, Paola Gongra, and Antonio Ribeiro. 2017. *Avaliação para Negócios de Impacto Social: Guia Prático*. São Paulo: Atermisia.
- Simonsen, Jesper, and Toni Robertson. 2013. *Routledge International Handbook of Participatory Design*. Routledge International Handbooks. New York, NY: Routledge.
- Sismondo, Sergio. 2010. *An Introduction to Science and Technology Studies*. Chichester: Wiley-Blackwell.
- Sivarajah, Uthayasankar, Vishanth J. P. Weerakkody, P. Waller, Habin Lee, Zahir Irani, Y. Choi, R. Morgan, and Y. Glikman. 2016. "The Role of E-Participation and

- Open Data in Evidence-Based Policy Decision Making in Local Government." <https://bradscholars.brad.ac.uk/handle/10454/14107>.
- Skipper, Yvonne, and Debra Pepler. 2021. "Knowledge Mobilization: Stepping into Interdependent and Relational Space Using Co-Creation." *Action Research* 19 (3): 588–605. <https://doi.org/10.1177/1476750320960810>.
- Slaughter, Sheila, and Gary Rhoades. 2000. "The Neo-liberal University." *New Labor Forum* 6: 73–79.
- Smith, Karl. 2012. "From Dividual and Individual Selves to Porous Subjects." *The Australian Journal of Anthropology* 23 (1): 50–64.
- Smith, Simon, Vicky Ward, and Allan House. 2011. "'Impact' in the Proposals for the UK's Research Excellence Framework: Shifting the Boundaries of Academic Autonomy." *Research Policy* 40 (10): 1369–1379.
- Sobrinho, Elena. "Environmental Justice and the Relational Politics of Abolition." Remote presentation, Annual Meeting of the Society for the Social Study of Science, Toronto, October 9, 2021.
- Souza, Siddharth de, and Siddhi Gupta. 2021. "Collective Brainstorming to Imagine a Feminist Labour Collective." *FemLab.Co* (blog), 2021. <https://femlab.co/2021/10/04/collective-brainstorming-to-imagine-a-feminist-labour-collective/>.
- Spears, Taylor Clancy. 2014. "Engineering Value, Engineering Risk: What Derivatives Quants Know and What Their Models Do." PhD Thesis, University of Edinburgh.
- Spiro, Lisa. 2012. "'This Is Why We Fight': Defining the Values of the Digital Humanities." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 16–35. Minneapolis, MN: University of Minnesota Press.
- Stanley, Liz, and Sue Wise. 1993. *Breaking Out Again: Feminist Ontology and Epistemology*. London: Routledge.
- Statistics Canada. 2023. "Census Profile, 2021 Census of Population." Statistics Canada Catalogue no. 98-316-X2021001. Ottawa. Released November 15, 2023. <https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E>.
- Steinbuch, Maarten. 2016. "Towards the 4th Generation University." Online: <https://steinbuch.wordpress.com/2016/07/23/towards-the-4th-generation-university/>.
- Stephan, Paula, Reinhilde Veugelers, and Jian Wang. 2017. "Reviewers Are Blinkered by Bibliometrics." *Nature* 544 (7651): 411–412.
- Stillman, Sarah. 2007. "'The Missing White Girl Syndrome': Disappeared Women and Media Activism." *Gender & Development* 15 (3): 491–502. <https://doi.org/10.1080/13552070701630665>.
- Stuart, Toby E. 2000. "Interorganizational Alliances and the Performance of Firms: A Study of Growth and Innovation Rates in a High-Technology Industry." *Strategic Management Journal* 21 (8): 791–811.

- Suárez Val, Helena. 2020. "Datos Discordantes. Información Pública Sobre Femicidio En Uruguay." *Mundos Plurales* 7 (1): 53–78.
- Surowiecki, James. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York, NY: Anchor Doubleday.
- Sutherland, Georgina et al. 2016. *Media Representations of Violence Against Women and Their Children: Key Findings and Future Directions*. Sydney: ANROWS. <https://www.anrows.org.au/publication/media-representations-of-violence-against-women-and-their-children-key-findings-and-future-directions/>.
- Tax Justice Network. 2015. *The Greatest Invention: Tax and the Campaign for a Just Society*. Margate: Commonwealth Publishing.
- The eQuality Project. "The eQuality Project." Accessed December 31, 2021. <https://www.equalityproject.ca/>.
- The eQuality Project. "Arts Exchange." Accessed December 31, 2021. <https://www.equalityproject.ca/resources/art-exchange/>.
- The eQuality Project. "Disconnection Challenge." Accessed December 31, 2021. <https://www.equalityproject.ca/resources/disconnectionchallenge/>.
- Toffler, Alvin. 1970. *Future Shock*. New York, NY: Random House.
- Topp, Stephanie M., Marta Schaaf, Veena Sriram, Kerry Scott, Sarah L. Dalglish, Erica Marie Nelson, Rajasulochana SR, et al. 2021. "Power Analysis in Health Policy and Systems Research: A Guide to Research Conceptualisation." *BMJ Global Health* 6 (11): e007268.
- Touraine, Alain. 1994. *Crítica de la modernidad*. Buenos Aires: Fondo de Cultura Económica.
- Triadafilopoulos, Triadafilos. 2021. "The Foundations, Limits, and Consequences of Immigration Exceptionalism in Canada." *American Review of Canadian Studies* 51 (1): 3–17.
- Turner, Victor. 1969. *The Ritual Process: Structure and Anti-Structure*. Chicago, IL: Aldine.
- Ucbasaran, Deniz, Paul Westhead, and Mike Wright. 2001. "The Focus of Entrepreneurial Research: Contextual and Process Issues." *Entrepreneurship Theory and Practice* 25 (4): 57–80.
- University of Ottawa. *Transformation 2030: Building the University of Tomorrow*. https://transformation2030.uottawa.ca/sites/default/files/DGC19_1135_T2030_BROCHURE_ENG_WEB4.pdf.
- UN OHCHR. 2015. "UN rights expert calls all States to establish a 'Femicide Watch'." Press release. Geneva: UN OHCHR. 24 November 2015. <https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=16796&LangID=E>.
- UN Women. 2020. "Infographic: The Shadow Pandemic – Violence Against Women and Girls and COVID-19". UN Women. 6 April 2020. <https://www.unwomen.org>.

- org/digital-library/multimedia/2020/4/infographic-covid19-violence-against-women-and-girls.
- UNDP, and FICCI. 2021. "Impact of COVID-19 and Industry 4.0 on Future of Work for Women." https://www.in.undp.org/content/india/en/home/library/womens_empowerment/Impact_of_COVID19_and_Industry_on_Future_of_Work_for_Women.html.
- UNESCO Institute for Lifelong Learning, Shanghai Open University. 2023. International Trends of Lifelong Learning in Higher Education. Research report. UNESCO. Online: <https://unesdoc.unesco.org/ark:/48223/pf0000385339>.
- Unicef, UN Women, UNFPA, ILO, OSRSG/VAG. 2013. "Breaking the Silence on Violence against Indigenous Girls, Adolescents and Young Women." <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2013/5/violence-against-indigenous-women-and-girls.pdf?la=en&vs=1457>.
- Ustek Spilda, Funda, Kelle Howson, Hannah Johnston, Alessio Bertolini, Patrick Feuerstein, Louise Bezuidenhout, Oğuz Alyanak, and Mark Graham. 2022. "Is Anonymity Dead?: Doing Critical Research on Digital Labour Platforms through Platform Interfaces." *Work Organisation, Labour & Globalisation* 16 (1): 72–87. <https://doi.org/10.13169/workorgalaboglob.16.1.0072>
- Utrecht University. 2020. *Strategic Plan 2025*. <https://www.uu.nl/en/organisation/strategic-plan-2025/the-strategic-plan>.
- Utrecht University. 2021. "Recognition and Rewards Vision." University policy. Utrecht: Utrecht University. <https://www.uu.nl/sites/default/files/UU-Recognition-and-Rewards-Vision.pdf>.
- Van der Sluijs, Jeroen. P. 2002. "A Way Out of the Credibility Crisis of Models Used in Integrated Environmental Assessment." *Futures* 34 (2): 133–146.
- Van der Zwaan, Bert. 2017. *Higher Education in 2040: A Global Approach*. Amsterdam: Amsterdam University Press.
- Van Dijck, José, and Thomas Poell. 2013. "Understanding Social Media Logic." *Media and Communication* 1 (1): 2–14.
- Van Dijck, José. 2014. "Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology." *Surveillance & Society* 12 (2): 197–208.
- Van Doorn, Niels. 2017. "Platform Labor: On the Gendered and Racialized Exploitation of Low-Income Service Work in the On-Demand Economy." *Information, Communication & Society* 20 (6): 898–914.
- Van Es, Karin, and Mirko Tobias Schäfer. 2017. "New Brave World." In *The Datafied Society. Studying Culture through Data*, edited by Mirko Tobias Schäfer, and Karin van Es, 13–22. Amsterdam: Amsterdam University Press.
- Van Es, Karin, and Mirko Tobias Schäfer. 2017. *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.

- Van Es, Karin, Maranke Wieringa and Mirko Tobias Schäfer. 2018. "Tool Criticism: From Digital Methods to Digital Methodology." ACM WS.2 2018 Proceedings of the 2nd International Conference on Web Studies. <https://doi.org/10.1145/3240431.3240436>.
- Van Es, Karin, Mirko Tobias Schäfer and Maranke Wieringa. 2021. "Tool Criticism and the Computational Turn: A 'Methodological Moment' in Media and Communication Studies." *M&K Medien & Kommunikationswissenschaft* 69 (1): 46-64.
- Van Schie, Gerwin, Irene Westra, and Mirko Tobias Schäfer. 2017. "Get Your Hands Dirty: Doing Ethical Research in a Big Data Era." In *The Datafied Society: Social Research in the Age of Big Data*, edited by Mirko Tobias Schäfer and Karin van Es, 183-200. Amsterdam: Amsterdam University Press.
- Veerbeek, Joris, Jeroen Bakker, Sahra Mohamed, Meilina Reksoprodjo, Tom Stienen, Mirko Tobias Schäfer. 2022. Online antisemitism [Online antisemitism]. Research report, Utrecht University. Online: <https://dataschool.nl/en/nieuws/antisemitisme-in-de-nederlandse-online-sfeer-in-2020/>.
- Veerbeek, Joris, Karin van Es and Eggo Müller. 2021. "Professional follow-up communication as an indicator of public value: Analyzing Dutch PSB's role in societal debates through topic modelling of press coverage." *Journal of Communication Inquiry*. <https://doi.org/10.1177/01968599211039213>.
- Veerbeek, Joris, Karin van Es and Eggo Müller. 2022. "Public Broadcasting and Topic Diversity in The Netherlands: Mentions of Public Broadcasters' Programming in Newspapers as Indicators of Pluralism." *Javnost*. <https://doi.org/10.1080/13183222.2022.2067956>.
- Venturini, Tommaso, Anders Munk, and Axel Meunier. 2018. "Data-Sprinting: A Public Approach to Digital Research." In *Handbook of Interdisciplinary Research Methods*, edited by Celia Lury, Rachel Fensham, Alexandra Heller-Nicholas, Sybille Lammes, Angela Last, Mike Michael, and Emma Uprichard. London: Routledge.
- Verran, Helen. 2021. "Narrating a Number and Staying With the Trouble of Value." In *The Data Journalism Handbook: Towards a Critical Data Practice*, edited by Liliana Bounegru and Jonathan Gray, 55-64. Amsterdam: Amsterdam University Press. <https://doi.org/10.2307/j.ctv1qr6smr.11>.
- Villasante, Tomás. 1999. "Cuatro redes para hacer transformaciones sustentables." *Política y Sociedad* 31: 37-54.
- Vliegthart, Rens and Liesbet van Zoonen. 2011. "Power to the Frame: Bringing Sociology Back to Frame Analysis." *European Journal of Communication*, 26 (2): 101-115. <https://doi.org/10.1177/0267323111404838>.
- Voorn, Eduard. 2017. "Professor Emile Aarts: 'Heb Als Ondernemer de Guts Om Big Data-Stappen Te Zetten.'" *De Ondernemer*, May 9, 2017. <https://www.deondernemer.nl/nieuwsbericht/171036/professor-emile-aarts-heb-als-ondernemer-de-guts-om-big-data-stappen-te-zetten>.

- VSNU. 2014. *Code of Conduct for Academic Practice: Principles of Good Academic Teaching and Research*. Translated by Metamorfose Vertalingen BV. The Hague: Association of Universities in The Netherlands. [http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_\(version2014\).pdf](http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code%20of_Conduct_for_Academic_Practice_2004_(version2014).pdf).
- Walter, Maggie, Tahu Kukutai, Stephanie Russo Carroll, and Desi Rodriguez-Lonebear. 2020. *Indigenous Data Sovereignty and Policy*, 1st edition. Abingdon: Routledge. <https://doi.org/10.4324/9780429273957>.
- Webster, Kim et al. 2018. *Australians' Attitudes to Violence Against Women and Gender Equality. Findings from the 2017 National Community Attitudes towards Violence against Women Survey (NCAS)*. Sydney: ANROWS. <https://ncas.anrows.org.au/wp-content/uploads/2018/11/NCAS-report-2018.pdf>.
- Weil, David, and Amanda Pyles. 2005. "Why Complain? Complaints, Compliance, and the Problem of Enforcement in the U.S. Workplace." *Comparative Labor Law & Policy Journal* 27 (1): 59–92.
- The White House. 2010. "Remarks by the President to the United Nations General Assembly." September 23, 2010. <https://obamawhitehouse.archives.gov/the-press-office/2010/09/23/remarks-president-united-nations-general-assembly>.
- Wilton, Robert. 2004. "Keeping Your Distance: Balancing Political Engagement and Scientific Autonomy with a Psychiatric Consumer/Survivor Group." In *Radical Theory, Critical Praxis: Making a Difference Beyond the Academy?*, edited by Duncan Fuller and Rob Kitchin, 116–131. Vernon and Victoria: Praxis (e)Press.
- Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019a. "Networked but Commodified: The (Dis)Embeddedness of Digital Labour in the Gig Economy." *Sociology* 53 (5): 931–950. <https://doi.org/10.1177/0038038519828906>.
- Wood, Alex J., Mark Graham, Vili Lehdonvirta, and Isis Hjorth. 2019b. "Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy." *Work, Employment and Society* 33 (1): 56–75. <https://doi.org/10.1177/0950017018785616>.
- Woodcock, Jamie, and Mark Graham. 2020. *The Gig Economy: A Critical Introduction*. Cambridge: Polity Press.
- Wright, Mike. 2007. *Academic Entrepreneurship in Europe*. Cheltenham: Edward Elgar Publishing.
- Young, Meg, Michael Katell, and P. M. Krafft. 2019. "Municipal Surveillance Regulation and Algorithmic Accountability." *Big Data & Society* 6 (2): 2053951719868492.
- Ziewitz, Malte. 2016. "Governing Algorithms: Myth, Mess, and Methods." *Science, Technology, & Human Values* 41 (1): 3–16. <https://doi.org/10.1177/0162243915608948>.
- Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. London: Profile Books.
- Zuckerman, Ethan. 2014. "New Media, New Civics?" *Policy & Internet* 6 (2): 151–168. <https://doi.org/10.1002/1944-2866.POI360>.
- Zwitter, Andrej. 2014. "Big Data Ethics." *Big Data & Society* 1 (2): 1–6.

Index

- academia 14, 19-22, 26, 28, 31, 35, 36, 51-56, 64, 77, 79, 90, 98, 99, 143, 144, 191, 231, 239, 244-5, 261, 266
- accountability 28, 37, 88, 135, 161, 167, 170-1, 261
- activism 36, 47, 134, 138, 197, 210
- advocacy 20, 36, 38-41, 105, 128, 136, 208, 210, 212, 253-255, 262, 265
- agency 42, 71, 107, 111-115, 171
- citizen agency 28, 111
- algorithmization 14, 62, 65
- America
 - Latin 29, 103, 104, 109, 112, 134, 135
 - United States of 88, 139, 206, 262
- Argentina 135, 139
- artificial intelligence (AI) 11, 20-1, 28, 30, 62, 66, 70-1, 88, 235, 236, 253
- autonomy 36, 55, 70

- bias 48-56, 137, 141-143, 16-6, 169, 227, 236, 238, 243, 245, 262-3
- Brazil 29, 88, 136, 192, 193, 195, 196, 202

- Canada 29, 88-9, 93, 94, 161, 164, 209, 252, 253, 257
- capital
 - human 23
 - financial 39, 40, 150
 - social 39, 197
 - symbolic 239
- change
 - societal 56, 80
 - technological 14, 73, 230
- citizen participation 53, 66, 223,
- co-design 124, 127, 134-141, 169-70, 172, 222
- colleges 54, 164; *see also* universities
- communities
 - diverse 20, 23, 25-6, 28, 76, 107, 111, 113, 162, 170, 208, 209, 244, 246
 - epistemic 240
 - local 36
 - immigrant 175
 - racialized 162-169
 - marginalized 49, 56, 185
 - of professionals 236-8, 243
 - of understanding 109
- critical data studies 21, 27-8, 37, 39-44

- data
 - analysis 28, 68, 70, 121-29, 165, 174, 180, 182-90, 206, 241, 261
 - audits 28, 111
 - collection 21, 72, 79, 91, 113, 122-5, 134-137, 154, 165-7, 173-5, 182, 184, 187, 206, 225
 - feminism 112, 136-9, 143, 206
 - infrastructure 30, 38-44, 187, 205, 207-212
 - literacy 22, 28, 29, 103-9, 114, 179, 187, 226, 228; *see also* digital literacy
 - open 29, 38, 41-2, 49, 74, 87-9, 93, 134, 179-80, 187, 206-9
 - personal 103, 106, 110, 166, 175
 - resources 21, 74
 - science 29, 38, 65, 68, 135, 144, 172-3, 175, 180, 182-3, 186
 - subjects 103-9, 113, 183,
 - work 20-31, 37, 112, 180, 182, 261
- datafication 14-5, 21-2, 25, 28, 29, 30, 61-81, 103-15, 163, 165-6, 219-21, 229-30
- democratization
 - of political systems 53
 - of science 53
- digitalization 20
- disaggregation 167
- discrimination 53, 154, 162-3, 168, 175, 253

- engagement
 - community 111, 115, 166, 170-2, 175, 255
 - stakeholder 66, 129
 - societal 19, 26, 68
 - public 20, 26
 - with platform companies 152, 155-5
- entrepreneurial research 28, 61-81
- entrepreneurship
 - academic 63
 - see also* research, entrepreneurial
- ethics
 - data 41, 68, 70, 72, 75, 78, 220, 22-4, 226
 - of entrepreneurial research 71-2
 - research 258
- ethnography 237, 265
- expertise
 - community 107
 - disciplinary 186
 - domain 69
 - interactional 239-41
 - scientific 49, 54
 - situated 108
 - of professionals 77
 - of stakeholders 23, 26

- femicide, *see also* violence, gendered 133-44
- fieldwork 30, 50, 53, 64, 66, 198, 235, 241, 244

- gig
 - economy 29, 110, 151, 15-6
 - workers 152, 153
- God trick (Haraway) 48, 53, 55
- governance 165, 167, 171, 220, 221, 223, 224
- government
 - open 28, 29, 87, 88, 89-99, 179, 180

- grants 25, 98, 155
grievance studies 51-2
- humanities 14, 19, 20, 24, 63, 64, 69, 70, 73, 91, 206, 253
- ideology 49, 52
imaginaries
 data 214
 technological 62
immigration 162, 164, 168, 174
impact
 agendas 23
 social 65, 68
 societal 27, 30, 43, 52, 80, 81, 90, 227, 244, 261
impartiality 43, 242
independence 27, 29, 36, 42, 43, 53, 55, 77, 78, 89, 155, 156
information asymmetries 29, 149, 152, 154, 156
integrity 30, 42, 71-79, 88, 108
interdisciplinary
 approaches 14, 266
 collaboration 21, 25, 64, 80, 136, 143, 186
 project 30, 73, 76, 123, 251
 research 23, 24, 26, 77, 129
interviews 41, 42, 67, 93, 137, 139, 141, 153-5, 174, 198, 200, 209, 228, 239
- justice 22, 29, 48, 56, 89, 111, 161, 175, 179-187, 202, 205-14
- knowledge
 conference 76, 224
 dissemination 76, 229
 economies 20
 primary source knowledge 139, 240
 production 23, 36, 37, 47-49, 55, 69, 103-105, 112, 152, 255
 scientific 48, 49
 society 48, 55
 tacit 65, 237, 239-241, 245
 transfer 22, 62, 67, 129, 222-224, 244
- labor
 academic 15, 24, 27, 30
 emotional 136
 invisible 67, 75
 practices 22, 29, 151-154, 157
- machine learning *see* artificial intelligence (AI)
marginalization 79, 162
media studies 222, 226
methodology
 qualitative 64, 266
 relational 105
metrics 25, 70, 90, 92, 96, 227
- Mexico 88, 135
modernity, reflexive 48, 55
- neoliberalism 19, 22, 79, 110
Netherlands 72, 80, 220, 221, 224, 244
networks 39, 40, 65, 75, 97, 99, 125, 196, 198, 201, 228, 229, 266
neutrality
 reflexive neutrality 27, 47, 55, 56
news media 125, 126, 128, 134, 142
NGOs 111, 154
- objectivity *see* impartiality
observation, participatory 23, 65, 79
open-source 142
opportunity 14, 62, 64, 80, 96-98, 113, 122, 137, 154, 181, 186, 213, 230
- participation
 civic *see* citizen participation
partners, external 15, 26, 62, 64-68, 70-74, 78, 79, 223, 266
peer review 49, 51, 91, 98
perspective
 feminist 51, 53; *see also* feminism
 insider 242, 245
- platform
 economy 29, 149, 151, 152, 156, 157
 society 49
platformization 20, 104, 105
positionality *see* situatedness
positivism 47
postmodernism 51-52, 54
 reflexive postmodernity 55
post-truth 50, 54, 56
power asymmetries 78, 226
practitioners *see* external partners
programming 150, 214, 263, 362
 programming language 241
promotion (career) 98-99, 152, 254, 255.
Puerto Rico 257
- racism 143, 161, 162, 165, 169-171, 175
regulation/s (legal) 88, 93, 109, 167, 184, 263
research
 action 23, 28, 29, 35, 36, 53, 61, 103-105, 108, 109, 112, 114, 133, 136, 148, 149, 152, 154-157, 256
 applied 20, 22, 28, 64, 87, 97-99
 collaborative 21-26, 30, 31, 63, 67, 73, 75, 90, 129, 219, 221, 229, 231, 261, 265, 267
 cooperative 262
 covert 243
 ethnographic 182
 entrepreneurial 28, 61-65, 68, 71, 76, 77, 78-81
 transdisciplinary 23, 24, 30, 64, 219, 225, 227-229, 231, 251

- scholarship
 - applied 89, 92, 93, 98
 - public 37
 - traditional 91
- science communication 20
- scientism 56
- sector
 - private 181, 184
 - public 40, 42, 121, 129, 130, 164, 173
- sensorialities 112
- situatedness 43, 48, 136, 266; *see also*
 - positionality
- smart city 41-42
- social media 37, 42, 54, 61, 68, 70-72, 121, 123, 125-129, 137, 144, 197, 200, 239, 252
- social sciences 14, 20, 24, 35, 36, 38, 50
- sociolinguistics 127
- South Africa 88, 152, 154
- stakeholders 14, 15, 19, 20, 23-27, 30, 38, 40, 41, 43, 44, 66, 88, 90, 121, 124, 152, 156, 157, 164, 170, 172, 175, 220, 222
- statistics 42, 162, 186, 237, 241
- subaltern 197
- subjectivity 103-107, 109-111, 114
- subjectivities 28, 103, 109, 113-115
- survey/s 42, 95, 123, 124, 127, 129, 139, 153, 154, 172, 175, 198
- tenure 92, 98, 99, 254, 255
- transdisciplinarity 266
- transparency 43, 87-89, 97, 142, 149, 151, 175, 185, 209, 211, 212, 220, 243, 243
 - fiscal 88, 89
- Twitter [now X] 68, 69, 25, 126, 128
- universities *see* colleges
- Uruguay 104, 110, 134, 139, 141-145
- values
 - progressive 53
 - public 15, 113
- violence
 - domestic 121, 122, 126
 - family 121-130
 - gendered 123, 124, 129; *see also* feminicide
- workshops 13, 30, 39, 41, 69, 76, 93, 99, 124, 129, 198, 206, 207, 208-214, 253
- youth 110, 191, 200, 202, 251-253, 256, 257

The influence of austerity measures and neoliberal ideologies has sparked discussions about the relevance and value of academic institutions, particularly in the humanities and social sciences. Universities are redirecting academic focus towards greater societal engagement. This book argues that academia has much to gain by moving beyond its institutional walls, in our case, by doing data work with stakeholders and civil society. This collaborative work benefits citizens in our democratic, open societies and advances our knowledge economies. *Collaborative Research in the Datafied Society* offers a combination of theoretical insights, practical methodologies, and case studies, showcasing the power of collaborative research with stakeholders across diverse communities and civil society to tackle challenges that address pressing issues stemming from data practices and social justice issues. Taken together, the book's chapters formulate relevant concepts for grounding societally engaged research in the theories and methodologies from different disciplines. In addition, the book informs university administrators and research directors how to advance academia effectively towards mutual knowledge transfer with societal sectors.

Mirko Tobias Schäfer is Associate Professor of AI, Data & Society at Utrecht University's research area 'Governing the Digital Society' and the Department for Information and Computing Sciences. He is co-founder and Sciences lead of the Data School.

Karin van Es is Associate Professor, Media and Culture Studies, and project lead Humanities at Data School at Utrecht University. Her research contributes to the fields of software studies and critical data studies.

Tracey P. Lauriault is Associate Professor, Critical Media and Big Data and board member of the Institute for Data Science at Carleton University in Canada.