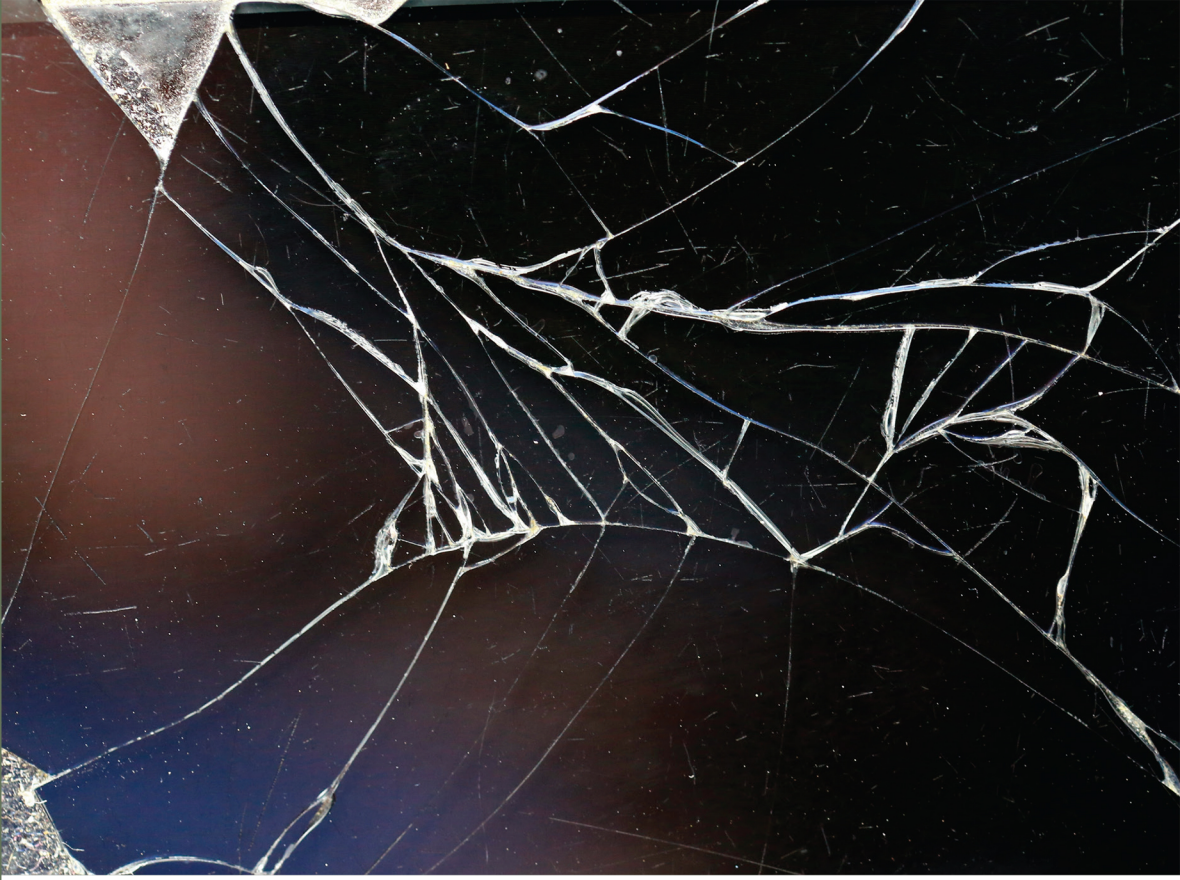

INVISIBLE SEARCH AND ONLINE SEARCH ENGINES

THE UBIQUITY OF SEARCH IN EVERYDAY LIFE

JUTTA HAIDER AND OLOF SUNDIN

ROUTLEDGE

INVISIBLE SEARCH AND ONLINE SEARCH ENGINES

Invisible Search and Online Search Engines considers the use of search engines in contemporary everyday life and the challenges this poses for media and information literacy. Looking for mediated information is mostly done online and arbitrated by the various tools and devices that people carry with them on a daily basis. Because of this, search engines have a significant impact on the structure of our lives, and personal and public memories. Haider and Sundin consider what this means for society, while also uniting research on information retrieval with research on how people actually look for and encounter information.

Search engines are now one of society's key infrastructures for knowing and becoming informed. While their use is dispersed across myriads of social practices, where they have acquired close to naturalised positions, they are commercially and technically centralised. Arguing that search, searching, and search engines have become so widely used that we have stopped noticing them, Haider and Sundin consider what it means to be so reliant on this all-encompassing and increasingly invisible information infrastructure.

Invisible Search and Online Search Engines is the first book to approach search and search engines from a perspective that combines insights from the technical expertise of information science research with a social science and humanities approach. As such, the book should be essential reading for academics, researchers, and students working on and studying information science, library and information science (LIS), media studies, journalism, digital cultures, and educational sciences.

Jutta Haider is an Associate Professor in Information Studies at the Department of Arts and Cultural Sciences, Lund University, Sweden. She has published widely on digital cultures' emerging conditions for production, use, and distribution of knowledge and information. This includes work on environmental information and on knowledge institutions, including encyclopaedias, search engines, and the scholarly communication system.

Olof Sundin is Professor in Information Studies at the Department of Arts and Cultural Sciences, Lund University, Sweden. His work concerns mainly information literacies and practices in various settings. More precisely, he researches the configuration of information in contemporary society, the construction of trustworthiness of public knowledge, and information searching and use.

"This is an excellent book that provides a unique perspective on the field of web search studies. Jutta Haider and Olof Sundin place search where it is best placed: in everyday life. Search has become neither the sole activity of finding reliable information on important topics nor just looking up facts or trivia but is used for both purposes interchangeably. This book, which is well-grounded in the literature from different fields, provides an excellent basis for understanding search as an integral part of everyday life."

Dirk Lewandowski, *Hamburg University of Applied Sciences, Germany*

"This is a marvelously written, beautifully researched volume which explores the centrality of search to our lives. It ranges fluently across a number of fields to produce fresh insights into topics such as the creation of new temporalities and meaning-making in information infrastructures."

Geoffrey Bowker, *University of California Irvine, USA*

"Search is probably the most important concept in library and information science and is increasing, as demonstrated in the book, and is also studied by many other disciplines, including media studies and science studies. The book does a great job by presenting new knowledge of such an interdisciplinary nature. It is well known that there are many perspectives (or 'paradigms') in these fields, and this book advocates that search and search engines should be based in the perspective of everyday life and practice studies, and thereby it represents an important new voice in this extremely important field."

Birger Hjørland, *University of Copenhagen, Denmark*

INVISIBLE SEARCH AND ONLINE SEARCH ENGINES

The Ubiquity of Search in Everyday Life

Jutta Haider and Olof Sundin

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2019
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
52 Vanderbilt Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Haider, Jutta, author. | Sundin, Olof, author.
Title: Invisible search and online search engines : the ubiquity of search in everyday life / Jutta Haider, Olof Sundin.
Description: Milton Park, Abingdon, Oxon ; New York : Routledge, 2019. | Includes bibliographical references and index.
Identifiers: LCCN 2018052680 (print) | LCCN 2019000340 (ebook) | ISBN 9780429448546 (ebook) | ISBN 9780429828010 (epub) | ISBN 9780429828003 (mobi/kindle) | ISBN 9780429828027 (web pdf) | ISBN 9781138328600 (hardback: alk. paper) | ISBN 9781138328617 (pbk. : alk. paper)
Subjects: LCSH: Internet searching. | Web search engines--Social aspects. | Information behavior.
Classification: LCC ZA4230 (ebook) | LCC ZA4230 .H35 2019 (print) | DDC 025.0425--dc23
LC record available at https://lccn.loc.gov/2018052680

ISBN: 978-1-138-32860-0 (hbk)
ISBN: 978-1-138-32861-7 (pbk)
ISBN: 978-0-429-44854-6 (ebk)

DOI: 10.4324/9780429448546

Typeset in Bembo
by Taylor & Francis Books

CONTENTS

ACKNOWLEDGMENTS

This book is the result of years of research involving three projects that have been carried out in collaboration with many people whose input we could not have been without, but who at the same time cannot be blamed for anything written in the book.

The book is primarily an outcome of the project "Knowledge in a Digital World: Trust, Credibility & Relevance on the Web" (The Swedish Research Council, 2012–05632). We would, therefore, first of all like to thank all our colleagues on this project. Particularly, we want to thank those colleagues who worked together with us in carrying out the focus group discussions that we use as illustrative examples throughout the book: Cecilia Andersson, Hanna Carlsson, and Sara Kjellberg. It has been a pleasure to work with you. Without our discussions and written communications, this book would never been produced. Having said that, we of course would also like to thank all participants in our focus groups who generously shared their experiences of and ideas about searching in everyday life. The inside knowledge of your daily routines has been crucial for our project. We would also like to address a special thank you to Erik J. Olsson, who co-led the project, and to his colleagues in the philosophy LUIQ research group. You have provided us with many valuable comments that have challenged us to develop our argumentation, at the same time as the uniqueness of our contribution from information science/studies has become clearer. To some extent, the book also relies on the findings from our projects "Open Data – The Materiality and Fragmentation of Facts" (Erik Philip Sörensen Foundation) as well as "Algorithms and Literacies: Young People's Understanding and Society's Expectations" (The Swedish Research Council, 2017–03631).

Earlier analyses of the focus group discussions we draw on in the book, as well as other related work, have been published by us in the following publications: "Invisible search: Information literacy in the Swedish Curriculum for Compulsory

Schools" in *Nordic Journal of Digital Literacy*, 2015, 10(4), 193–209; "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices" in *Journal of Documentation*, 2016, 72(6), 990–1007 (together with Hanna Carlsson); "The structuring of information through search: Sorting waste with Google" in *Aslib Journal of Information Management*, 2016, 68(4), 390–406; "Searching for delegated knowledge in elementary schools" in Proceedings of ISIC, the Information Behaviour Conference, Zadar, Croatia, 20–23 September, 2016: Part 2. Information Research, 2017, 22(1), paper isic1618 (together with Hanna Carlsson), "Researchers' online visibility: Tensions of visibility, trust and reputation", in press at *Online Information Review* (together with Sara Kjellberg). In particular, thoughts developed in two publications have found their way into the book, where they were further developed and deepened, namely: "The search-ification of everyday life and the mundane-ification of search" in *Journal of Documentation*, 2017, 73(2), 224–243 (together with Cecilia Andersson, Hanna Carlsson, and Sara Kjellberg), and "Controlling the urge to search: Studying the informational texture of practices by exploring the missing element" in *Information Research*, 2017, 22(1), CoLIS paper 1613.

Drafts of individual chapters or parts of chapters have been presented at various conferences and seminars. The panel (organised by Olof Sundin and Trine Schreiber) "Is there a need for re-claiming search in information literacy practices research?" at CoLIS9 in Uppsala 2016 provided us with an early opportunity to test one of the leading hypotheses of the book – namely that the practice turn in information behaviour and literacy research has made search and search engines less visible. We are grateful to Cecilia Andersson, Andrew Cox, Helena Francke, Camilla Moring, Ola Pilerot, and Trine Schreiber, who all contributed to this thought-provoking debate then and on other occasions. This theme appears in particular in Chapters Two and Six of the book. The paper "Controlling the urge to search: Studying the informational texture of practices by exploring the missing element" was presented at the same conference by Jutta Haider, and ideas from this paper have been developed and interwoven particularly into Chapter Four. The conference paper "Critical algorithm literacies: An emerging framework" was presented at the ECREA Digital Culture and Communication Section Conference in Brighton in the autumn of 2016. The conference paper "Facts, fake and information literacy: A conceptual discussion of critical evaluation of information" was presented at ISIC 2018 in Krakow the autumn of 2018. These two papers, both presented by Olof Sundin, have primarily been developed into Chapter Five of this book as have ideas tested by Jutta Haider at a research seminar at The Swedish School of Library and Information Science, University of Borås. Thank you to Jan Nolin, Louise Limberg, and Annemaree Lloyd and the other seminar participants for perceptive and encouraging comments. Ideas that found their way into Chapters Three and Four have also been presented at the Department of Science and Technology Studies, Alpen-Adria-Universität in Klagenfurt, Austria, by Jutta Haider. Thank you to Daniel Barben and Nils Matzner for commenting and asking thought-provoking input. In sum, the feedback and many constructive comments we were

given on all these occasions have been very important and we appreciate and are grateful for the valuable input of all participants.

The Department of Arts and Cultural Sciences at Lund University and all its staff provides an excellent environment for conducting research. Chapter One was presented at our Information Studies Research Seminar. Our thinking owes depth to the collective endeavour of our own research group, Information Practices: Communication, Culture and Society and we would, therefore, also like to mention, in addition to the names presented in the introduction of the acknowledgement, Lisa Dahlquist, Charlotte Hagström, Fredrik Hanell, Linnea Lindsköld, Nora Schmidt, Björn Magnusson Staaf, Fredrik Åström, and Carin Graminius (thank you for your help with translating from Chinese). In particular we need to thank Johanna Rivano Eckerdal for her intellectual companionship and generosity. Furthermore, in our meetings with students on the Master's Programme in Archival Studies, Library and Information Studies, and Museology (ALM) and in the bachelor program in Digital Cultures we on a daily basis develop our thinking through lectures, supervision, and other dialogues. The library facilities at our faculty are first-class, and we would especially like to mention Ulrika Karlsson who, in her capacity as subject librarian for information studies, has demonstrated what library support beyond facilitating access can be. We are also grateful to our editor Heidi Lowther and editorial assistant Elizabeth Risch for their support and responsiveness.

We have had the opportunity to discuss our research with wider society on a number of occasions, with journalists, teachers, and librarians around Sweden, and through our commissioned work at the Swedish National Agency for Education, the Swedish Schools Inspectorate, The Internet Foundation in Sweden and Government Offices of Sweden. Our outreach contacts have certainly made us better prepared to communicate the societal relevance of our research, which we hope shines through in the book.

Finally, but certainly not least, we would like to express our gratitude to the invited readers who we asked to comment on separate chapters or the entire manuscript. Dirk Lewandowski, Isto Huvila, Pamela McKenzie, and Nils Pharo have read one to two chapters and provided us with invaluable comments that helped us develop our argument. A final and special thank you goes to Ola Pilerot and Trine Schreiber who, despite a tight time frame, read through the whole manuscript. Your generous and insightful comments late in the process have developed the manuscript into something much better than it would have been without you.

Enjoy,
Jutta Haider & Olof Sundin
Lund, Sweden, Autumn 2018.

1

INTRODUCTION

This is a book about search. It is concerned with how search, searching, and with them search engines have become so widely used that we have stopped noticing them. It is thus more accurately a book about invisible search. One of society's key infrastructures for knowing and becoming informed is computerised systems supporting the search for and locating of documents and information. The use of these systems, search engines, is curiously dispersed and centralised at the same time. It is dispersed across a vast array of social practices in which it has acquired close to naturalised positions (Hillis et al. 2013), while it is commercially and technically centralised and controlled by a handful of very dominant companies, especially one extremely powerful global player, Google. In the course of this, we can call it double movement, search has all but vanished from sight. Invisibility is often highlighted as one of the key features of an infrastructure (Star 1999) and clearly, in this sense it is safe to say that search engines are a fundamental information infrastructure. Yet, what does that mean more specifically? How do people deal with search engines? How do we research their use and which strands of previous research help us understand this all-encompassing, increasingly invisible information infrastructure?

With digital tools now permeating most aspects of society, the use of search engines has become integral to everyday life on many levels. Searching for recipes or the bus timetable, for medical advice, for old neighbourhood photos or the weather report, searching for work-related documents, legal guidance or for the whereabouts of former colleagues, all this is done instantaneously on the same interface and using the same devices, thus inscribing itself into various parts of life. Search permeates myriads of social practices and everyday life at all levels, but it often remains invisible. It appears to be simple and is done effortlessly. Yet, this effortless simplicity with which online search intersects with everyday life in so many different situations conceals an astounding complexity. Accordingly, various strands of research traditions have for a long time been dealing with different

DOI: 10.4324/9780429448546-1

aspects of search and search engines. In fact, an entire set of considerations comes to the fore relating to issues such as the ways in which social practices are integrated with technology, with understanding what information might be and do in a certain situation, how to make sense of it in a specific context in relation to search technology, or in which ways to trust or question it. Furthermore, being searchable is today not only often seen as a feature of information, but information is also moulded to fit the shape provided by the tools used for searching for it, and, more often than not, this is a web search engine (Gillespie 2017; Haider 2014; Kallinikos et al. 2010). Inversely, this also means that information that is not produced in conformity with the rules laid down by dominant search engines gets buried and is made less visible (Mulligan & Griffin 2018, pp. 569–570). Ultimately this – we can call it *search-ification* – of everyday life relates to the ways in which an increasingly invisible information infrastructure is entangled across culture and its practices and to what means we have at our disposal for understanding and making sense of these entanglements (see also Sundin et al. 2017).

To look for, find, or retrieve information has always been one of the central concerns of libraries, as for information science, and at least since the 1940s and 1950s information science has been concerned with search in computerised systems more specifically. In a narrow sense and when the focus is on technical systems, the field is called information retrieval. In a broader sense and with a focus on people, it is often referred to as information behaviour. More colloquially, talk is of (online) searching or searching for information. Accordingly, these days society's most important information retrieval systems are simply called search engines. Looking for mediated information is mostly done online and arbitrated by the various tools and devices that people carry with them on a daily basis. In addition, various algorithms and not least economic interests organise search. This way, search engines contribute to structuring private as much as professional lives and public and personal memories in ways that might not be immediately obvious.

This *search-ification* of everyday life is also connected to the fact that contemporary general-purpose web search engines are by most measures easy to use and return in most cases results that are experienced as useful. What used to be complicated-to-use information retrieval systems, integrated with expensive (for users) databases designed for specific professional groups and frequently queried by information professionals or librarians, have for many turned into an unassuming query box or are integrated into a mobile phone by voice recognition via virtual personal assistants. Full-text searches across an enormous and ever-growing index are carried out at a speed that makes them appear to be almost instantaneous. Neither professional education, knowledge of query languages, nor controlled vocabularies are needed to query a general-purpose web search engine and to do so in a way that is adequate for most purposes. No subscriptions are required to use them. In fact, they are increasingly seen to replace specialised retrieval systems or these have begun to emulate general-purpose search engines in order to appear user-friendly. As using search engines and doing so proficiently became feasible for the general public, searching became not only inserted into all kinds of social

practices, but was also de-professionalised. We propose to call this the *mundane-ification* of search (Sundin et al. 2017).

Using search engines is now just another ordinary activity, hard to discern and increasingly difficult to study and also teach. This is not least due to the way that the technical workings of the system have become increasingly opaque to users the simpler to use a search engine appears. This is amplified by the fact that control over this crucial information infrastructure is in the hands of commercial organisations whose business model is based on their having full control over their index (database), their algorithms and their user data and whose "core unit of exchange" is traffic (van Couvering 2008, p. 177).

Everyday life and social practices

This book is framed around the notion of everyday life. A focus on everyday life is quite common in order to subsume all kinds of practices, behaviours, and periods that are not related either to education or the workplace. Obviously, this can pertain to extremely different situations and thus play out differently when related to search. Everyday life is a fundamentally temporal notion (Adam 1995). Its emergence and meaning is closely connected to industrialisation and to the structuring of production in capitalist society and in particular of the welfare state (Nowotny 1994). Everything is tightly structured, negotiated, and controlled by the rules of the market and of work, in terms of holidays, protected working time per week or day, shifts, overlay time, weekends, and so on. The advent of the notion of everyday life is also part of a shift from a focus on production to one on consumption, which makes possible and also requires non-working time conceptualised as leisure time or quality time and importantly also as time for consumption. Only when we think of time as something that can be subject to commodification, which it is in capitalist society, does the concept of everyday life make sense. "Everyday life has become the bracket combining work and so-called free time; the private 'spending of time' and the public spending form a new combination in everyday life", writes Helga Nowotny (1994, p. 103) and she continues: "The great public institutions of the state and the economy, and their temporal perspectives, are confronted with the temporal perspectives of the citizens and employees, the economic subjects". In information science, time is an undertheorised notion and Reijo Savolainen called already in 2006 (p. 124) for the "need to develop a research agenda in order to approach the temporal issues of information seeking more systematically". Search engines are one of today's most important information and communication technologies and clearly their enormous presence has implications not only for how we think *of* time but also for how we think, experience, and practice time.

Talking of different temporal perspectives that meet and collide, opens up for an understanding of everyday life that is more attuned to the intermingling of different structures of control, of various allegiances and demands than to a predominantly chronological notion, where one period follows after the other.

Rather, the illusive concept of everyday life can be understood as something that is part of all life in different ways and not necessarily clearly distinguished from that which is "not everyday life". We can think of it as reaching into situations and periods of, for instance, work or education or participation in civic life. Helga Nowotny (1994) talks of the increasing blurring of the boundaries between public and private time in the phenomenon of everyday life. Interestingly for our investigation, she also identifies a related change in human perceptions of time that she discusses in relation to the ubiquity and dominance of communication technology in contemporary society. While work is sometimes conceptually distinct from everyday life as its negative and defining anti-thesis, it is also part of people's everyday life in a more common-sense understanding. Interestingly, for many, today it is precisely the use of ICTs and general-purpose search engines that in part ties work and non-work together (see also Rosa 2015). This could be considered in terms of a context collapse, where the same tool is used in different situations, e.g. work and private life or family and friends, thus intermingling not just expectations and ways of doing things, but also time scales and temporal perspectives.

Everyday life also denotes *the normal*, that which is not exceptional. In this sense, everyday life is connected to habits, to repetition, all deeply temporal notions. Our account of online search in everyday life foregrounds precisely this perception of online search as the usual, the normal, as an activity that has established itself as a central part of various routine practices, that is, as one of the taken-for-granted elements making up the practice in question. Social practices are generally theorised in terms of arrangements of activities that connect in specific ways and which in these specific arrangements are observable across groups, that is, they are neither individual nor exceptional. Considerable attention, empirical and conceptual, has been paid to the intricate relationships between information and social practices. We will return to some of this work in Chapter Two. However, less attention has been given to the way in which information technology is implicated in re-arranging practices and what this means for how we engage with and organise everyday life. Chuck Moran (2015, p. 299), drawing on the foundational work of Barbara Adams (e.g. 1995), maintains "[s]ocial practices, particularly those with significant technological components, are organizing temporality in exciting ways". And he continues: "The passing of time is becoming less important than the accrual of possible events or of sequences that can be activated whenever. Indeed, these practices challenge the very meaning of temporal terms such as memory, event, and even the contemporary". They also challenge us to rethink our understanding of routines in social practices and of how and when they are constituted. Hartmut Rosa (2015, p. 235) draws our attention to what he calls a "paradigm shift" regarding everyday time strategies and practices. He sees the classic modern everyday routine conceptually superseded by a "temporalized everyday time of late modernity"; the first one dominated by schedules and predefined time, the latter by flexibility and deadlines. Specifically mobile technology, Rosa convincingly develops, and its potential to support the negotiation of flexibility in everyday life is implicated in propelling such a late modern time regime, where "[t]ime

orderings are /…/ individually and flexibly created within time itself" (Rosa 2015, p. 235). However subtle, search engines constitute a considerable technological component, to draw on Moran's expression, in an ever-increasing number of social practices and their ubiquity and dominance are irrevocably tied to mobile technology. Relatedly, the role of mediated information in everyday life is being reorganised on just about every level possible and the way in which search engines are implicated in this is what is at issue in this book.

Between systems and users, search engines and search

Web search engines are formed by their use. They are not static systems that always perform in exactly the same way; rather, they are supposed to adapt to different users' needs or expectations and also as they are being used to train and improve themselves constantly. This happens on two – closely interlinked – levels, an individual level and a collective level. On the one hand, web search engines are personalised systems. They adapt search results to individual users, or actually more precisely to their user profiles. Mostly this is done in order to increase the chances of returning search results that are perceived as being relevant and useful by the person searching and ultimately this helps to improve the placing of adverts. Increasingly and tied to the growing importance of machine learning and artificial intelligence (AI), search engines, and specifically Google, have started to develop into "suggest engines", where active searching is not even necessary as the system is supposed to anticipate what a user wants or needs without a search having to be carried out, at least not in the sense of a person entering a query.

On the other hand, at a collective level, the aggregated queries and user logs are used to improve the system, also with the help of machine learning. Knowing common spelling mistakes, understanding which terms are often or rarely entered together, or which links are followed helps making reasonable assumptions about what others also might want to find (or be suggested). This is also shown to the users in the form of suggestions for further, related searches or through the autocomplete function, where search terms are completed as they are being typed. While this clearly has the effect of helping to formulate relevant searches and thus supports the users, there is also another side to it. Displaying common searches further re-enforces their very dominance, thus contributing to further bolstering their significance and cultural meaning. It is a kind of *Matthew Effect*, where the rich become richer or, in our case, the known becomes more known. In certain cases, this way – given the trust vested in search engines as neutral brokers of facts – truth and factuality can be established. As Helga Nowotny poignantly (2015, p. 58) notes, "[w]hen Google sought to gauge what people were thinking, it became what people were thinking". The relationships between users, documents, and systems or between searchers, information, and engines are extremely intimate, in many ways they are constitutive of each other, and as search engines and search – i.e. what we do with them – become increasingly invisible, it becomes increasingly obvious how agency here is not something reserved for humans, but a function

also inherent in technology – a function distributed and in constant flux within and between different actors (Latour 2005).

Yet, despite the growing importance of search engines and their dominant role in how information is accessed and organised in society, and despite the long-standing concern of the discipline with information retrieval and information behaviour, the contributions of information science to understanding search and search engines are scattered across different subfields. Research on systems occurs on one side and research on users most often on the other. Both traditions thrive, yet, despite attempts (e.g. Ingwersen & Järvelin 2005), they do not always connect. This is to the detriment of research on web search engines and their use in what is commonly called everyday life. One reason for this might lie with the different theoretical starting points of the two fields. The traditions simply see different things and they see things differently.

One strong strand of research is concerned with information systems: How they can be built, evaluated, and optimised and, within limits, how they are used. The literature and research produced in this area has a focus on technical issues. Accordingly, it is institutionally often situated in the universities' various technical departments. Research in this field does consider behavioural aspects of search, that is how systems are used, yet these issues are mainly addressed in experimental laboratory studies and very rarely in everyday life situations (Jansen & Rieh 2010). On the other hand, and especially since the 1980s, there has been a reaction to the perceived dominance of this technical orientation in the field, and user studies emerged as a further strong area in information science. Here, the focus is on people and their various encounters with information in more natural, everyday life settings. Often, these are studied by drawing on general theories developed in the wider social sciences and in psychology (for a compilation, see Case & Given 2016). Here, the focus is on information behaviours and needs, cognitive factors, tasks, and literacies, and most recently – in line with the so-called practice turn that has occurred in most social sciences – information practices. Admittedly, there are sometimes vast theoretical and methodological divides and more or less subtle differences between the areas united in this user perspective. Yet, for all intents and purposes they are on "speaking terms". What unites them is their interest in the use of information or engagements with information aside from dedicated information systems. These are positioned as just one factor amongst many, which people engage with in ways that can be captured as informational, and over the years information systems have moved further and further away from the primary locus of interest. Curiously, this has happened at the same time as dedicated information systems, which search engines after all are, have moved into ordinary everyday life with a force that has made them more or less ubiquitous. In Chapter Two, we will focus in greater detail on earlier information science research that has a bearing on the formative role of search engines in society and will present it alongside research from other related disciplines in an attempt to bring them together.

For now, as not least Ralph Schroeder (2015, 2018) has emphasised, while information science has contributed with foundational principles for how search engines work, as a discipline it has not yet contributed in any significant way to the

understanding of searching and search in everyday life. Partly this can be explained by the divided nature of the discipline we have just discussed where relevant research is scattered across different subfields that are not often connected. However, search engines remind us how these two strands of research depend on each other and are intimately connected. We cannot study search without saying something about search engines and this cannot be done without acknowledging the foundational contributions made by the technical side of information science to how they function. Their technical functioning is no coincidence but is rooted in certain understandings of what they should deliver in terms of results and how these are achieved and presented to people. Conversely, we cannot study search engines without considering search, that is, how they are used. This cannot be done without acknowledging the important contributions of research to how information is looked for, stumbled upon or otherwise embedded in everyday life and its practices. As Wanda Orlikowski (2007, p. 1445) puts it,

> we see how the researcher's Google search is constituted by the performativity of computers, networks, softwares, algorithms, directories, databases, and infrastructure, as these are enacted by the human agencies entailed in their design, construction and operation.

Together these actors form a "sociomaterial assemblage" (ibid.). We suggest that a sociomaterial theoretical understanding of search, which will be presented further in Chapter Three, has potential for bringing different research fields in information science together. In this sense, this book is an attempt at moving the study of information systems back into the centre of information science, yet fully acknowledging and continuing the work that has been carried out on the situational and material character of information and specifically on information in social practices.

How is a search result relevant?

In order to be meaningful, a search for information needs to lead to relevant results. It is thus not surprising that relevance is a central concept for developing and understanding information retrieval systems, which search engines after all are. "Historically", Tefko Saracevic (2016, p. 11) writes "relevance actually crept in unannounced", but it has since developed into one of the most important in information science. "The concept of relevance is the key issue to the functioning and evaluation of IR systems", writes, for instance, Pia Borlund (2003, p. 923), pinpointing what is at stake. To date, most search engines return a list of links to web pages with information related to the search request. At the most rudimentary level, they look for whether the word searched for appears on a website or not or whether the pattern of an image corresponds to images online. Applying semantics to web data, search engines increasingly also return what is more properly described as answers to specific questions directly on the search results page. Either way, the results have to be displayed in an order that makes sense to the person who carried out the search in the first place. This order is established according to certain relevance criteria. Quite simply, what comes first in the list should be more

relevant than what follows. It is primarily due to their capacity to deliver relevant results and show them high up in the results list that a search engine becomes useful and popular in the first place. Hence, getting the order in which search results are presented right is crucial, not least for the commercial success of a search engine. The notion of relevance however, while it sounds straightforward is an old and returning problem in information science. This is not surprising given that one of the discipline's oldest concerns is about developing tools for locating information and information resources. In 1975, Tefko Saracevic (1975, p. 324) wrote, in the first paper of a series of literature reviews on relevance research, that "[t]he success of information science, whatever there is, is due to the fact that it did address itself to relevance". As Saracevic reminds us, relevance has been tweaked and researched on since the 1930s in the writings of Samuel C. Bradford and in the tradition of information retrieval since the 1940s (Saracevic 1975).

Simplified, when relevance is discussed in information science, once again the two strands of research, that is a focus on systems and a focus on users, as briefly presented above, are advanced as frames of reference and put in opposition to each other (e.g. Cosijn 2010). The systems approach regards relevance as a "property of the relation between the content of the document and the system's search criteria" (Cosijn 2010, p. 4512). This approach to relevance has been developed in the information retrieval tradition and in tests on controlled document collections (e.g. the Cranfield experiments in the 1960s and later in the Text REtrieval Conference (TREC)). It is dependent on the information system and criteria set up by experts to establish whether a document that was returned for a search is relevant or not. It is this in a binary way, either something is relevant or it is not. Here, in particular two notions have emerged as central, namely the conceptual couple *precision* and *recall*. They can be traced back to the mid-1950s, but are still of great significance (Saracevic 2016). In simple terms, recall measures the number of retrieved relevant documents in relation to all relevant documents in a collection, while precision measures the number of retrieved relevant documents in relation to the total number of retrieved documents (also non-relevant documents). The higher the share of retrieved relevant documents, the higher a system's so-called retrieval effectiveness is said to be. If all retrieved documents are relevant, which is the ideal outcome, precision is said to be 1 (see also Lewandowski 2018). In relation to general-purpose web search engines, measuring precision is feasible and also done in various studies, for instance to compare the retrieval effectiveness of different search engines (Lewandowski 2011). In contrast, calculating recall is more or less impossible, since there is no way of establishing what is in the collection in the first place.

In the user tradition, on the other hand, relevance is seen as it is understood by the user. Here, relevance is not binary, but graded. It is defined by how it is experienced, as a relation between information and user. Often, the dynamic character of relevance is emphasised. Accordingly, there is an interest in how relevance evolves and changes during a process. Altogether, it should be kept in mind that relevance is a fundamentally relative concept. It is deeply situated and depends on human judgement. In other words, it is dependent on people considering something to be meaningful and useful in specific situations. Birger Hjørland (2010)

challenges the received view of the dichotomy between a system approach and a user approach to understanding relevance. Spinning on an example taken from the field of medicine, which fits well with our own example with vaccines, Hjørland (2010, p. 226) suggests "the user normally does not have the qualifications to evaluate his or her information need". He also emphasises that the so-called system research approach to relevance in fact is contingent on experts assessing the topical relevance prior to tests. That is, it is not the system that ranks the relevance of a certain document, but experts and only subsequently is the system assessed based on its capacity to retrieve the documents designated as relevant in response to certain queries.

Nevertheless, relevance is, as Michel Buckland (2017, p. 161) reminds us "subjective, idiosyncratic, hard to predict and unstable". Paradoxically this malleability might be the precise reason for its success and persistent relevance in the field (see also Nolin 2009). Also, relevance is a concept that is not only multidimensional (Borlund 2003) and hard to delineate theoretically, but whose meaning also slips away and changes as the relationship between the found information or retrieved document, the system used for searching, and the user changes in situations of application. Hence, the concept of pertinence has been introduced to better capture the aspect of usefulness for the individual users in their specific situation. As Anthony C. Foskett (1996, p. 16) explains;

> a document retrieved in answer to a request may be *useful* to the enquirer, but its utility may change; for example, if we retrieve the same document in a second search, it will have lost its utility the second time around. Its *relevance* will not have changed, but the enquirer's view of it will.

In other words, pertinence is about utility and usefulness. Thus, while "relevance is a consensus judgement, pertinence is an individual judgement" (ibid.). What is relevant to one person in one situation might change over time due to the changed understanding of a topic. Depending on the situation, the same person can regard the same piece of information or the same document as returned in response to the exact same query as very relevant or entirely irrelevant. Conversely, in similar situations different people might judge different pieces of information to be relevant or irrelevant, depending on their circumstances.

What these different approaches, emerging from a user study tradition, to conceptualising relevance have in common is a focus on the individual. Yet, what the scholarly literature in information science rarely relates relevance to are societal or cultural concerns (for exceptions, see Hjørland 2010; Saracevic 2016; Sundin & Johannisson 2005). However, what is relevant and even useful information for an individual might not be relevant for society at large and vice versa. What is often referred to as topical or subject relevance (e.g. Saracevic 2016) is close to what we could call societal relevance. One difference though, while the notion of topical or subject relevance is developed and mostly given meaning in relation to books and journal articles, societal relevance also includes other aspects of the relation between information and society. When taking the question of relevance out of the traditional information science context of controlled collections to general web

search engines we argue that the notion of topical or subject relevance is too limited in its scope. Consider the example of vaccination: An individual who is sceptical of the common scientific understanding of vaccination assesses the relevance of websites about vaccination very differently from someone who believes in the standard scientific understanding of vaccination, its benefits and the potential risks involved. Accordingly, in all likelihood they will also judge the relevance of search results returned in response to their queries differently. Yet, at a societal level, there is a scientific consensus that vaccinations cannot be positioned as a solely individual choice. Many vaccination programmes are only effective if a level of so-called herd immunity is reached and this can only be achieved if most of the population are vaccinated. Clearly, in addition to relevance in the sense of "is this an adequate result for this search?" and pertinence, meaning "is this useful for me?" (user), there is a further layer, which shapes relevance and pertinence, but which also has a complicated life of its own, societal relevance, or, perhaps more accurately, societal interest. Control – algorithmic, political, discursive, and affective control – over societal interest and its constantly shifting relation to pertinence, document-centred, and individual relevance is what is at issue in debates about function and shape of public knowledge that is expressed in talk of the disappearing trust in established institutions, of the dismissal of expertise and the challenges to social cohesion. Undoubtedly, users' assessments of relevance are of importance, yet these needs to be related to different types of societal interests and how this relating happens can be ground for conflicts and tensions. We refer to this potential tension, between individual assessments of relevance and societal aspects of it as a *friction of relevance*, a concept we put into play in Chapter Four. One of the most important contributions from information science to understanding relevance as being fundamentally societal might somewhat unexpectedly come from another strand of research, namely bibliometrics.

In information science, bibliometrics is typically considered as belonging to the field of scholarly communication. That is, the quantitative study of the artefacts of scholarly communication. In fact, Eugene Garfield, who developed a number of the most foundational bibliometric and scientometric principles, indirectly also contributed to the foundation of relevance assessment in search engines (de Bellis 2009, pp. 285–288; Halavais 2018[2009]). At a basic level, when developing the journal impact factor (JIF), Garfield postulated that, if the articles in a journal receive a high number of citations over a certain period of time, then this journal can be considered as having a higher impact than a journal whose articles are cited less frequently over the same period. The same assumptions have since been applied to books, authors, countries, or institutions and so on. Citations establish relationships between these entities. In turn, this establishes a form of relative relevance in relation to the community doing the citing and receiving the references. Translated to the web, this means that links to a website can be seen as an expression of relevance. In very simple terms, the more links point to a website, the higher it can be assumed is its relevance to the community doing the linking and also regarding the topic it addresses. In this way, popularity becomes a way to establish relevance. This principle has famously been introduced and perfected by

Google. It is expressed in its so-called PageRank algorithm, named after one of Google's founders, Larry Page (Page et al. 1999). PageRank exploits the network of links in the same way that bibliometric principles, developed by Eugene Garfield in the 1960s, exploit references and citations as connections between documents so as to establish the relative importance of scholarly journals and articles and later also authors and topics. A difference worth mentioning between the PageRank algorithm and the workings of a citation index is the editorial and peer-review system that, despite all its faults and weaknesses, functions as some kind of gatekeeper, balancing – at least to some degree – the effect of ranking by popularity. As Peter Ingwersen and Kalervo Järvelin (2005, p. 238) point out when discussing the relation between bibliometrics and the functioning of PageRank and similar algorithms: "Citations (or inlinks) given to the objects are also manifestations of socio-cognitive relevance judgements made by people over time". In addition, in order to further increase relevance, contemporary search engines often personalise and adapt results according to the data they have on the person (or actually the user profile) using them. Taken together, popularity and personalisation (ideally at least) help to balance relevance as something that lies between individual utility and societal arrangements – yet always calibrated in relation to the most basic principle contemporary web search engines operate within, and that is, of course, the market and commercial success in this market.

The elephant in the room: Google

There is no way around acknowledging that, what Ken Hillis and his colleagues (2013, p. 1) aptly describe as an "astonishing naturalization of the process of search in everyday life" and what we choose to call the concurrent *mundaneification* of search, happened in tandem with the rise of one specific commercial search engine, Google. Each minute more than three million queries are submitted to Google Search. Globally in the second decade of the 2000s the market share of Google Search constantly lies between 85% and 90% (StatCounter n.d. a). Notable exceptions are Russia and especially China (StatCounter n.d.b). In both countries other large commercial search engines exist – with their own indexes and search tools: Baidu in China and Yandex in Russia. This has, of course, to do with censorship and related issues of information control, especially in China, but also with character sets used to write in these country's languages. However, both Yandex' and Baidu's interfaces and ways of displaying results are noticeably inspired by Google. All available statistics point in one direction and that is that Google dominates all aspects of the search engine market, largest share of searches on all devices, most visitors to its site, most users of its browser, Chrome, which in effect is a search engine, and so on. This data is, of course, time sensitive, but, in order to understand the role of search engines as a social and cultural phenomenon Google has to be taken seriously and has to be situated in its specific historic context.

Google's rise started in 1998 in the midst of a period of commercialisation of the internet, the so-called dot.com boom and then bust (van Couvering 2007). An uncluttered interface, different from the search portals that dominated at the time, and the above-mentioned algorithm for ordering lists with search results distinguished Google from its competitors. As mentioned above, Google's then newly patented algorithm, PageRank, stipulated that links to web pages can be likened to votes indicating the relative importance of a web page (Mayer 2009). Put simply, a web page that receives more so-called in-links is considered more important than one receiving fewer of those in-links. In turn, the links from more important pages, its out-links, are given more weight than links coming from less important web pages. By exploiting the web's link structure in this way, Google could provide searchers with rankings of search results that were experienced as and considered to be vastly more relevant and pertinent than those of other search engines around at the time. As Elisabeth van Couvering (2007) points out, neither the technical solution nor the business idea were the sole factors behind Google's success. Google entered the market at a time when a change occurred in how search engines created value. It was part of a fundamental shift in the search engine business:

> No longer would the *audience* (the traditional media commodity sold to advertisers) be at the core of the search business. Now, the online commodity of choice would be *traffic* or the flow of visitors from one Website to another. /…/ [As] traffic emerged as a key commodity in its own right, sites which had as much traffic as possible – that is to say, as many people coming and going as possible – became the nexus of economic traffic.
>
> *van Couvering 2008, p. 196*

Since the invention and patenting of PageRank ensured Google's success, the search engine's algorithms have been improved and added to innumerable times to include new ranking signals, aiming at improving relevance, such as when or where a page was published, what a user searched for before, which device someone is using, and hundreds more (Lewandowski 2012). Also, while game changing when it was first introduced, in the two decades that followed, the importance of the PageRank algorithm gradually declined, not least of course because in-links as a measure of relevance can easily be gamed through, for instance, artificial link farming. Yet, despite the fact that no single web search engine dominates global society's information flows as completely as Google does, it is unknown how exactly its algorithms work. There is a serious lack of transparency and accountability that Google has in common with other globally dominant information companies, such as Amazon and Facebook. Yet, there is one industry with special interest in understanding how Google's algorithms function, the field of Search Engine Optimisation (SEO) or Search Engine Marketing (SEM). They have developed skills in interpreting and making sense of the bits and pieces that are made publicly available by Google in various publications, guidelines, tutorials, or

interviews and bring these together with what they themselves can establish through experiments and trial and error approaches. "Constant changes of the search algorithm make the relationship between Google and SEO companies highly dynamic, antagonistic and fluid", writes Pasko Bilić (2016, p. 7), and he continues: "Google monetises its unique visitors and information search. The connection between user intentions and useful search results is a moving target tangled in a web of socio-technical and economic decisions".

One place to look for evidence of what Google's algorithms actually do is the SEO/SEM sector's various industry publications. Reading industry publications, newsletters, and forums shows a curious hunt that is taking place where the SEO/SEM professionals "run" after the various animals after which Google chooses to name its algorithms and changes in them. In 2013 it was the hummingbird that was released "into the wild" and it was up to users and the professionals working with moulding information to be searchable in Google to establish what this hummingbird was doing. What kind of an animal was this algorithm or, probably more precisely, this algorithmic assemblage? The hummingbird, who followed in the path of first the panda and then the penguin, signalled a move into the direction of facilitating conversational searches. These are searches that are expressed like real-life questions in a natural language rather than simply query terms. Clearly, this is critical in order to meaningfully support voice search and other forms of voice interaction with personal assistants. It is quite non-transparent how Google's various algorithms and signals relate to each other, and what is part of what. Nevertheless, importantly, as voice-operation gains ground, there is a move towards providing answers directly rather than merely lists with links to where the answers can be found. This requires an understanding of a search in its proper context and not just as a task checking for the convergence of search terms. Google – like other search engines – approaches this challenge with semantic web technology and various artificial intelligence solutions. For instance, RankBrain is now part of the hummingbird algorithm. RankBrain helps to deduce the meaning of entirely new searches. Searches that have never been entered into the engine before, in order to be related to previous searches, require to be interpreted in a way that is based on understanding the meaning of terms in specific situations far beyond and very different from term matching. Importantly, RankBrain constantly improves itself; it learns and gets better at making its interpretations; it understands better the longer it is in use. Relating searches to other searches in this way together with voice-operation and the further merging of search into the background are developments that are hugely consequential for the way in which society's access to information is structured. Specifically, it has implications for the possibilities that remain for critical evaluation of information and of the tools organising access to information and more precisely for where to find an opening in the increasingly impenetrable surface we are presented with, into which to insert suitable instruments of critique.

Its successful algorithms are clearly not the only reason behind Google Search's popularity and its enormous commercial success. Google owns by far the largest index of the web, i.e. the continuously updated databases containing the downloaded webpages on

which their search software performs its searches. Through their many other services and businesses – now held by the holding company Alphabet – they also collect myriads of data from other sources not least by tracing people's actions and interactions online. Together, this amounts to a vast assemblage of continuously growing volumes of data whose real potential and risks for predictive analyses or the development of artificial intelligence applications will only become fully apparent in the future. In Chapter Three, we will consider in more detail how the fundamental notions of algorithm, index, and data are constituted and cast them in a theoretical understanding foregrounding their various constitutive entanglements with each other and their uses.

Google's market dominance is also connected to the development of other freely available tools, such as its cloud storage, text editing tools, e-mail, calendar, digitised books, news, and other, frequently pre-installed applications on smart phones and tablets, and of the Android platform. Being visible in and accessible through Google is of increasing importance for what we are able to know in the first place and how we are able to know – as individuals and as a society. The implications of this can hardly be overstated. The convergence of Google's different parts into one interlinked information eco-system in most parts of what is called the western world has implications for almost all information-related activities, directly mediated ones, but also others. Google's significance in society is often described in no humble terms and rightly so. For instance, Alex Halavais (2018[2009], pp. 6–7) calls Google a "touchstone of digital culture", while Ken Hillis and his colleagues talk of *Google as Culture* in the book of the same name and describe it as having attained a "socially consecrated status" (Hillis et al. 2013, p. 7). Google has also been theoretically framed, drawing on Michel Callon's concept, as an "obligatory passage point" (Mager 2009; Sundin & Carlsson 2016) and Siva Vaidhyanathan (2011) even talks of the *Googlization of Everything* as he draws attention to the profound implications of Google and its various platforms for shaping the very norms of digital communication. A further interesting concept in this regard is the *Google-effect* (Sparrow et al. 2011). It describes a presumed change in how we organise our memory "to include the computer and online search engines as an external memory system that can be accessed at will" (ibid., p. 778). This more or less conscious incorporation of the search engine into people's and in a way into society's collective external memory chimes well with Google's ambition to facilitate conversational searches and thus to gradually dissolve into society's background noise in order to function as a suggest rather than a search engine. What the SEO/SEM people are hunting becomes increasingly unclear, as they first have to identify the very species they are supposed to catch, yet while they try to guess which taxonomy to use for this task, the creature has morphed into yet another unknown species.

Searching for content versus searching for documents

Up until now we have made no clear distinction between information and document when referring to the results of a search. However, this hides an important dimension of the difference between the two, and one which is in the process of

dissolving which has implications for the increasing invisibility of search. The distinction can be exemplified by talking of searching for facts versus searching for documents containing the facts. Taking the example of the question about the height of Mount Everest, Brian Vickery famously describes it as follows: "'The height of Mt. Everest is given on page 900, volume 8, of Encyclopaedia Britannica'; in the other case, it would be 'The height of Mt. Everest is 29,002ft'" (Vickery 1961, pp. 2–3). In the first case, the answer is provided through a reference to the literature. In the second case, the answer is given by providing the fact as such. If we were to carry out the same search in Google, we would get both at the same time. The height of Mount Everest would be delivered right there on the search results page together with a link to Wikipedia and in fact millions of other links to all kinds of information about Mount Everest in various kinds of media and genres, from film, photos, documentaries, fiction films, video clips, private memories, journalistic content, and so on a personalised search results page, a great deal of it in the language of the country I am searching from. In fact, just entering the words "height of" into the query box already brings up a suggestion to look for Mount Everest's height, presumably because this is an extremely popular search. At the end of the first search results page, a series of suggestions for related searches will be made, also these presumably popular search queries that others have made right after or before searching for the height of Mount Everest.

This very basic distinction that Vickery was able to make between content and document and that seems so common-sensical and helpful when illustrated by means of a printed Encyclopaedia, becomes almost impossible to maintain in the case of Google, or for that matter other general-purpose search engines. Search engines merge the two types. They started in the tradition of helping to locate documents (web pages); however, since the search interface is the same interface on which the document is then displayed, from a user perspective this crucial distinction is hard to make out. This is even more difficult when search engines start to provide the information itself – as Google does – that is when, alongside the links, they also display answers right there on the search results page. This is the case with the so-called Google *Knowledge Graph*, which brings together open data sources, and increasingly its various e-business ventures (travel, shopping, etc.). Also, the featured snippets function, where the text from popular pages is pulled into the search results page and made visible right there before the well-known list starts, works to a similar effect. All these use technically different solutions, although from a user perspective the result is very similar: The search engine provides answers and not merely pointers to documents containing the answers. The enormous success of full text search combined with improved relevance rankings from the late 1990s onward was a big step in blurring this distinction, not necessarily on a theoretical level nor from the point of view of developers, but certainly for the average user and their everyday lives. With the growing importance of virtual assistants, artificial intelligence, and voice recognition, the distinction will soon be entirely counter-intuitive for most people. Google changes its role as an index and provider of links, framed as unbiased, to very openly becoming an arbiter and even a producer of knowledge presented as factual answers to questions, and becoming increasingly invisible in the process.

Thus, from the perspective of users, the search engine distorts the distinction between content and document and merges different document types into one. Effectively, the search engine results page can itself be conceptualised as a document, algorithmically generated, malleable, singular, and with no clear boundaries, yet a document nonetheless. At the same time, it is and has traditionally always been a document whose purpose is to help users locate other documents, like a card in the old-fashioned card catalogue would have. The search engine results page does this, but it also provides some of the epistemic content from the very documents it links to and it does this increasingly on the same interface and in the same devices (browser, computer, smartphone, and so one). This changes the status and purpose of the search engine results page as a document. In the process, the actual engine, the technicalities of search, become almost entirely opaque. This is achieved by combining full text search with semantic web technology. The content of the documents is marked up and thus described in ways that enable the search algorithms to determine what type of content they contain and what they *mean* and not just on which server they can be found. In this way Google returns not only a list of pages with information on the height of Mount Everest, but directly states 8848 metres high on its first results page. Katrine Juel Vang (2013, p. 258) sums up what is at stake:

> The Knowledge Graph represents a significant change in the role of Google. Until now, Google has simply disseminated links /…/ representing possibilities, allowing the users to choose themselves which of the results to examine further and which to dismiss. Now Google itself represents answers to the queries – instead of simply introducing the user to a range of possible sources, Google itself is the source.

Knowledge Graph and the use of semantic web technology and of learning algorithms to make it possible are, in fact, not unlike Paul Otlet's utopian vision of a Universal Book of Knowledge and his monographic principle (van den Heuvel 2008). The very idea of searching for pieces of information in a huge universe of documents in which films, letters, articles, photos, or books are piled together is often discussed in relation to the conceptual thinking of Paul Otlet (1868–1964), one of the founders of the so-called *documentalist* movement. Already in the early 20th century, Otlet aspired to organise and make findable single facts in various media rather than entire documents. He envisioned that

> [Information] will be recorded on separate leaves or cards rather than confined in volumes … By gathering these leaves together, and classifying and organizing them according to the headings of a reliable and detailed classification, we will create the "Universal Book of Knowledge".
>
> *Otlet 1903 in Rayward 2008, p. 15*

Google is not – of course – in any way the realisation of this Universal Book of Knowledge. It is, after all, a corporate venture which is still primarily about

matching consumers to adverts, and the size of its index is matched by the amount of surveillance data it holds on its users in order to perfect this exercise. However, the sheer scale of its index, together with the use of semantic web technology and the ensuing convergence of document, information, and infrastructure, bears witness to a similar ambition. In relation to Google's Knowledge Graph, there is another resemblance with Otlet's work. Where Otlet's vision was dependent on low-paid women, Knowledge Graph is dependent on Wikipedia and its free labour. In fact, most data presented in the Knowledge Graph function so far seems to be coming from Wikipedia.

Chapter conclusion: What is this book about and what is it not about?

This book wants to put the spotlight on the narratives, ideologies, technologies, ethical dimensions, and most of all the mundane practices tied to online search – its meaning, function, implications and limits – in contemporary society, and to give some indication as to how to study it. Information in its material form is organised in order to be found (or hidden, but that is a different – albeit related – story). Yet, this is not a book about how to effectively find information or how to organise it for this purpose. Rather, it is a book about how to use some of the rich research traditions in information science – sometimes referred to as information studies or library and information science – together with research from other related fields – from media and cultural studies to science and technology studies (STS), organisational studies and sociology – so as to understand what search does to everyday life and what everyday life does to search in contemporary society. The purpose is to carve out the roles search plays in what we do and how we know – from accessing bus timetables, to medical advice, to help with homework, educational or work issues, or trivial pastimes such as finding out what people we went to school with are up to today. As such, we are interested in the general-purpose search engine as a cultural artefact as it enters formative relation with people's practices in contemporary society. We are not concerned with how search behaviour could be optimised, nor are we are out after studying information seeking or searching generally. While search engines are the focal point of the book, we are less interested in detailing their exact technical workings, albeit of course that we want to highlight important fundamentals, than in how they are constituted in use and how they constitute our interactions with each other and with other engines. Web search engines do not exist on their own and nor are they the same all the time. They have changed considerably since their inception in the 1990s and will continue doing so. They are shaped by how they are used and in relation to other tools such as social media feeds or Wikipedia, various smart phone applications or the like, but also analogue tools and institutions such as archives or libraries, the media, government, civil society organisations, or educational institutions, businesses, the advertising sector, and also propaganda efforts. Most importantly, they are becoming increasingly invisible, not less formative, but harder and harder to

see. In a way, this book is an attempt to bring together research – our own and others – on these engines before they are entirely submerged and rendered invisible in society's basic infrastructures.

Chapter organisation

Invisible Search and Online Search Engines: The Ubiquity of Search in Everyday Life is organised into seven chapters. Following this first introductory chapter:

Chapter Two aims at familiarising readers with important currents in earlier research and to highlight positions and findings to build on. The chapter surveys some of the research traditions of information retrieval and of information behaviour, with the purpose of presenting a basic understanding and deepening some of the concepts introduced in Chapter One, but also broaches work from other fields.

Chapter Three presents and discusses a number of conceptual devices relevant for understanding the dominance and constitutive embeddedness of search and search engines in society. Three materially and technically consequential as well as theoretically and historically rich concepts – infrastructure, algorithms, data – are used to explicate some of the underlying technical structure of search engines and to situate search in a framework of sociomateriality.

Chapter Four reflects on some specific ways in which search is part of everyday life and its social practices. It is a largely empirical chapter, structured along the lines of thematic areas carved out in order to highlight some areas of significance for how search is embedded in society. Among others, the concept of friction of relevance is proposed to understand the relation between individual experiences and societal interests. Attention is also paid to not searching and to the role of search engines in confirming biases.

Chapter Five approaches search from the perspective of literacy and more specifically of media and information literacy by highlighting challenges for educational settings. The concept of infrastructural meaning making is suggested in order to shed light on the significance of sociotechnical systems that are highly trusted in society for filtering and providing information. In relation to this the limits of critical evaluation of information and of media and information literacy are discussed.

Chapter Six brings the issue of method to the fore. One underlying current in this chapter is the question of how to study a phenomenon and an activity that the researcher is deeply implicated in. The chapter has two parts, addressing complementary questions: How to investigate search and searching and how to use search engines for research?

The concluding chapter, Chapter Seven sums up the book's most important points and central concepts. The chapter concludes by relating the most salient points raised to a conceptual discussion attentive to issues of time and temporality. The book concludes by developing the analogy introduced in Chapter One, where Google is called the "elephant in the room", to ask "Maybe Google is not the elephant in the room, but the room itself?"

References

Adam, B. (1995). *Timewatch: The Social Analysis of Time*. Cambridge: Polity.

Buckland, M. (2017). *Information and society*. Cambridge, MA.: The MIT Press.

Bilić, P. (2016). "Search algorithms, hidden labour and information control". *Big Data & Society*, 3(1), 1–9.

Borlund, P. (2003). "The concept of relevance in IR." *Journal of the American Society for information Science and Technology*, 54(10), 913–925.

Case, D. & Given, L. (2016). *Looking for information: A survey of research on information seeking, needs, and behavior* (4th ed.). Bingley: Emerald.

Cosijn, E. (2010). "Relevance judgements and measurements." In M. Bates & M. N. Maack (Eds), *Encyclopedia of Library and Information Science* (3rd ed.). (pp. 4512–4519). New York: Taylor and Frances.

De Bellis, N. (2009). *Bibliometrics and citation analysis: From Science Citation Index to Cybermetrics*. Lanham, Md.: Scarecrow Press.

Foskett, A.C. (1996). *The subject approach to information* (5. ed.). London: Library Association Publ.

Gillespie, T. (2017). "Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem." *Information Communication and Society*, 20(1).

Haider, J. (2014). "Taking the environment online: Issue and link networks surrounding personal green living blogs. "*Online Information Review*, 38, 248–264.

Halavais, A. (2018[2009]). *Search engine society* (2nd ed.). Cambridge: Polity.

Hillis, K., Petit, M. & Jarrett, K. (2013). *Google and the culture of search*. New York: Routledge.

Hjørland, B. (2010). "The foundation of the concept of relevance." *Journal of the American Society for Information Science and Technology*, 61(2), 217–237.

Ingwersen, P. & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dortrecht: Springer.

Jansen, B. J., & Rieh, S. Y. (2010). "The seventeen theoretical constructs of information searching and information retrieval." *Journal of the American Society for Information Science and Technology*, 61(8), 1517–1534.

Juel Vang, K. (2013). "Ethics of Google's Knowledge Graph: Some considerations." *Journal of Information, Communication and Ethics in Society*, 11(4), 245–260.

Kallinikos, J., Aaltonen, A. and Marton, A. (2010). "A theory of digital objects." *First Monday*, 15(6). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3033/2564.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.

Lewandowski, D. (2011). "Evaluierung von Suchmaschinen." In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen 2: Neue Entwicklungen in der Web-Suche* (pp. 203–228). Heidelberg: Akademische Verlagsgesellschaft.

Lewandowski, D. (2012). "New perspectives on web search engine research." In D. Lewandowski (Ed.), *Web search engine research* (pp .1–16). Bingley, U.K.: Emerald.

Lewandowski, D. (2018). *Suchmaschinen verstehen* (2nd ed). Berlin, Heidelberg: Springer.

Mager, A. (2009). "Mediated health: Sociotechnical practices of providing and using online health information." *New Media & Society*, 11(7), 1123–1142.

Mayer, K. (2009). "On the sociometry of search engines. A historical review of methods." In K. Becker & F. Stalder (Eds), *Deep Search: The politics of search beyond Google*. Innsbruck: Studien Verlag, 54–72.

Moran, C. (2015). "Time as a social practice." *Time & Society*, 24(3), 283–303.

Mulligan, D. K. & Griffin, D. S. (2018). "Rescripting search to respect the right to truth." *Georgetown Law Technology Review*, 2(2), 557–584.

Nolin, J. (2009). "'Relevance' as a boundary concept: Reconsidering early information retrieval." *Journal of Documentation*, 65(5), 745–767.

Nowotny, H. (1994). *Time: The modern and postmodern experience*. Cambridge: Polity

Nowotny, H. (2015). *The cunning of uncertainty*. Cambridge & Malden: Polity.

Orlikowski, W. J. (2007). "Sociomaterial practices: Exploring technology at work." *Organization Studies*, 28(9), 1435–1448.

Page, L., Brin, S., Motwani, R. & Winograd, T. (1999). "The PageRank citation ranking: Bringing order to the web." Technical Report. Stanford InfoLab. Retrieved from http://ilpubs.stanford.edu:8090/422/.

Rayward, W. B. (2008). "European modernism and the information society: Introduction." In W. B. Rayward (Ed.), *European modernism and the information society: Informing the present, understanding the past* (pp. 1–26). Aldershot: Ashgate.

Rosa, H. (2015). *Social Acceleration: A new theory of modernity*. New York: Columbia University Press.

Saracevic, T. (1975). "Relevance: A review of and a framework for the thinking on the notion in information science." *Journal of the American Society for Information Science*, 26(6), 321–343.

Saracevic, T. (2016). "The notion of relevance in information science: Everybody knows what relevance is. But, what is it?" Synthesis lectures on information concepts, retrieval, and services. San Rafael, Calif.: Morgan & Claypool Publishers.

Savolainen, R. (2006). "Time as a context of information seeking." *Library and Information Science Research*, 28(1), 110–127.

Schroeder, R. (2015). "Does Google shape what we know?" *Prometheus*, 32(2), 145–160.

Schroeder, R. (2018). *Social theory after the internet: Media, technology, and globalization*. London: UCL Press.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). "Google effects on memory: Cognitive consequences of having information at our fingertips." *Science*, 333(6043), 776–778.

Star, S. L. (1999). "The ethnography of infrastructure." *American Behavioral Scientist*, 43(3), 377–391.

StatCounter. (n.d.a). "Worldwide desktop market share of leading search engines from January 2010 to July 2018." Statista – The Statistics Portal. Retrieved from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/.

StatCounter. (n.d.b). "Share of desktop search traffic originating from Google in selected countries as of June 2018." Statista – The Statistics Portal. Retrieved from https://www.statista.com/statistics/220534/googles-share-of-search-market-in-selected-countries/.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices." *Journal of Documentation*, 72(6), 990–1007.

Sundin, O., Haider, J., Andersson, C.Carlsson, H. & Kjellberg, S. (2017). "The search-ification of everyday life and the mundane-ification of search." *Journal of Documentation*, 73 (2), 224–243.

Sundin, O. & Johannisson, J. (2005). "The instrumentality of information needs and relevance." In F. Crestani & I. Ruthwen (Eds), *Context: Nature, Impact, and Role: 5th International Conference on Conceptions of Library and Information Sciences, CoLIS 2005, Glasgow, UK, 4–8 June 2005. Proceedings (pp. 107–118)*. Berlin, Heidelberg: Springer.

Vaidhyanathan, S. (2011). *The googlization of everything: (and why we should worry)*. Berkeley: University of California Press.

Van Couvering, E. (2007). "Is relevance relevant? Market, science, and war: Discourses of search engine quality." *Journal of Computer-Mediated Communication*, 12(3), 866–887.

Van Couvering, E. (2008). "The history of the Internet search engine: Navigational media and the traffic commodity." In A. Spink & M. Zimmer (Eds), *Web search: Information science and knowledge management, 14* (pp.177–206). Berlin, Heidelberg: Springer.

Van den Heuvel, C. (2008). "Building society, constructing knowledge, weaving the Web: Otlet's visualizations of a global information society and his concept of a universal civilization." In W. B. Rayward (Ed.), *European modernism and the information society: Informing the present, understanding the past* (pp. 127–154). Aldershot: Ashgate.

Vickery, B. C. (1961). *On retrieval system theory*. London: Butterworths.

2

PERSPECTIVES ON SEARCH

Close to ubiquitous access to search engines has reduced much of the variation in how people go about finding information. For a while, the term "Google generation" was floated (e.g. Nicholas et al. 2011; Rowlands et al. 2008) to account for this streamlining. However, it is not only about how to find information in a narrow sense, and the book title *Search Engine Society* (Halavais 2018[2009]) probably captures more accurately what is at stake, as does Siva Vaidhynathan's already mentioned famous dictum of the "Googlization of Everything" (2011) or Geert Lovink's notion of the "Society of the Query" (Lovink 2009). The profound changes in the way we search and find information transform our culture and re-organise our everyday life. Having said that, research on search and the use of search engines builds on a tradition that goes further back than web search engines. In this chapter, we paint in broad strokes a background picture by presenting a selection of earlier research and trace some of the traditions of significance for our purpose. From information science contribute *information behaviour* (IB) research and *information retrieval* (IR) research. These two fields bring an understanding of the interaction between the searcher and retrieval system (IR), as well as an understanding of people as information-seeking species in general, not necessarily connected to retrieval systems (IB). *Web search studies*, research on *information practices* and on *information seeking in everyday life* are to some extent outgrowths from information behaviour and information retrieval research. Finally, we draw sustenance from research on the *social shaping* and *the politics of search engines*, often published in the fields of organisational studies, science technology studies (STS) and media and communication studies. To start with, we take a look back in the research history of searching. Together, the chapter provides a rich tapestry of earlier research. Yet, despite the length of the chapter, it constitutes only a small selection. Anything else would simply be impossible.

DOI: 10.4324/9780429448546-2

A prelude

The use of tools to organise and find mediated information is a practice at least as old as the clay tablets making up the first known Sumerian libraries in Mesopotamia from around 3000 BC. To organise and retrieve information has always been closely tied to the material form of information and the techniques developed from that. Robert Darnton (2009) summarises the different technical developments for writing throughout history:

> from writing to the codex, 4,300 years; from the codex to movable type, 1,150 years; from movable type to the internet, 524 years; from the Internet to search engines, nineteen years; from search engines to Google's algorithmic relevance ranking, seven years.
>
> *p. 33*

To see the inventions in relation to information from a bird's eye view like this is almost dizzying. Clay tablets, the alphabet, the scroll, indexes, the codex, movable type, classification systems, parchment, papyrus, paper, the computer, digital retrieval systems, algorithms, and so forth all bring about radical changes to how information is communicated and searched for in society and, thus, also what we regard literacy in a broad understanding to be. This is not a chapter covering the history of searching. Still, this very short historical glimpse makes for a backdrop against which we can better understand the importance of the materiality of search.

The way to find information through published documents, organised according to specified principles, has been referred to as the *library paradigm* for information search (Trias i Mansilla & de la Rosa i Esteva 2013). The library paradigm stands in contrast to the *village paradigm* for finding information (ibid). The village paradigm builds on an older tradition of oral information culture where the only way to find out something, aside from your own experiences, was to ask someone who knows. While this simple dichotomy has shortcomings and glosses over various complexities, the archetypes drawn on can help to differentiate between the two types of finding information that also dominate contemporary online information practices. These days, when encountering something we do not know, besides of course asking people in person in our literal or metaphorical village or consulting the actual, physical (village) library, we can either use a tool for accessing documents or we can use social media or dedicated web services in order to ask more or less identifiable other people online. In this book, search is in focus and thus our interest lies predominantly in the library paradigm of accessing information. Nevertheless, the computerisation and algorithmisation of these archetypes has also meant that the ways of finding and accessing information that they represent have changed, as has the relationship between them. As we will see in Chapter Four, finding online information is, of course, not something solely done via search engines. Situated information, provided, for example, by questions and answers among "friends" on Facebook or other contacts in social media plays an important role in situations where trust in the source is crucial. To a large degree these two

archetypes are about where trust is predominantly, but of course not solely, located, that is either in people or in institutions. Thus, viewing for instance voice search and the increasing use of artificial intelligence (AI) assistants through the lens they provide, highlights how these developments are to a large extent about shifting the locus of trust from people and institutions to a technology that aims at merging and relocating it.

We will turn our attention back to search as it is understood in the library paradigm. Historically, bibliography – one of the predecessors of search engines – was developed as a method of listing and describing books in order to provide access to and an overview of what there is to know. Bibliographies were in the 1960s starting to be converted into bibliographical databases that were computerised collections of references to books, articles, reports, and so forth. Bibliography, as well as the tradition of encyclopaedism, is a tradition rooted in the humanities, but still an instrumental tradition that tries to bring order and access to the knowledge claims in the world. To search by using some kind of retrieval mechanism, such as the library or the bibliography, is, of course, nothing new, but, as we will show particularly in Chapter Four, activities of search have with contemporary technologies of search been integrated as parts of our practices of everyday life (Andersen 2018; Hillis et al. 2013). What once took hours, sometimes even days of travelling to a library or ploughing through archives, can now be accessed immediately. Importantly, as not least Tefko Saracevic highlights (2016, p. 10), this way the notion of relevance took center stage, alongside the aboutness traditionally in focus in bibliographic description.

The ambition to collect all representations of knowledge, more or less seen as given by God, in the world, has been around for a long time (Yeo 2001, p. 2). This desire unites bibliographers and encyclopaedists and, during the late 19th century, it was shared by idealist reformers – such as H. G. Wells and Paul Otlet – who strongly believed in the power of knowledge to bring peace and prosperity to the world. In his book *World Brain*, Wells writes: "There will not be an illiterate left in the world. There will hardly be an uninformed or misinformed person. And the brain of the whole mental network will be the Permanent World Brain" (Wells 1936–1938, in Yeo 2001, p. 2). The *World Brain* could be seen as one of many intellectual predecessors of Google (Rayward 2008). Another utopian, mentioned already in Chapter One, Paul Otlet, did not want to just produce a bibliography of everything published in the world. He goes further than that. Through his "monographic principle" he seems to anticipate the semantic web and a way to store and retrieve parts of texts rather than whole documents (Frohmann 2008). In the 20th century, this was followed, closer to the information science of today, by people like Vannevar Bush and his brain child *Memex* and Ted Nelson who conceptualised much of what is today known as hypertext and was thus foundational for the topography of the web.

In bibliography as well as in web search, an interest in the material side of information – the techniques for crawling, storing, indexing, searching, and so forth – goes hand in hand with ideas, visions and theories of knowledge representation. As discussed in Chapter One, the way web search is about to change even more through the increasing success of semantic web technology and artificial intelligence solutions, will make search less akin to using a bibliographic database,

where the guiding principle has been to provide searchers with a list of links in some kind of relevance order, and more like accessing a database of facts (or fact-like statements) where the sources for the individual pieces of information are more or less invisible to the users (Juel Vang 2013; Waller 2016). When iPhone users ask Siri which pizza restaurant to recommend, what kind of a search is that? Could search facilitated by the semantic web and technology like featured snippets be seen as a way in which the library and the village paradigms collude even more profoundly? We will get back to these questions in Chapter Three.

Search between information retrieval and information behaviour

Information behaviour (IB) research and information retrieval (IR) research make up two strong research traditions in information science, sometimes also reaching into neighbouring disciplines. Together, they have provided us with plenty of research to help us understand search and searching. However, the publishing patterns of the two traditions, with articles and conference papers constituting the predominant part, make it difficult to sum up the main findings. To provide a full literature review of the two traditions simply cannot be achieved in one volume. The fourth edition of *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior* (Case & Given 2016) consists of over 500 pages densely detailing just the information behaviour research tradition. Yet, the book still only captures a selection of the empirical findings that exist. One of the difficulties when describing the information behaviour tradition is what Tom Wilson describes as follows: "[t]he theoretical diversity is, perhaps, healthy, but the hope of theoretical conversion and unanimity has not been achieved" (2010, p. 31). In contrast to research on information behaviour, which is theoretically quite divided, the information retrieval tradition has a clearer paradigm with an agreed (although not always explicit) theoretical foundation, and established models and concepts. Information retrieval researchers have delivered a number of comprehensive books attempting to summarise the tradition's starting point and main models, as well as its findings (e.g. Ingwersen & Järvelin 2005; White 2016).

Research on information retrieval is in most historical accounts traced back to at least the 1950s. However, it is not until the 1960s that we can talk about an established information retrieval research tradition, with Gerard Salton's work at Cornell with *SMART Information Retrieval System* as one of the starting points (Saracevic 2016). The most important difference between early information retrieval and the type of modern information retrieval we are used to today is, as Amit Singhal (2001) summarises, the possibility to rank documents according to their usefulness in response to a query. This is fundamental considering the often enormous size of the text corpora that are available. The still existing Text REtrieval Conference (TREC), launched in the early 1990s by the US National Institute of Standards and which used people as judges to establish the usefulness of retrieved documents, was particularly important for developing methods for dealing with large text collections where ranking of results is critical (vector space

model, probabilistic model, inference network model) (Singhal 2001). Web search engines, however, also take advantage of the web's link structure. Here it is interesting to bring in Eugene Garfield's research on bibliometrics. His research has resulted in, among other things, the Science Citation Index (SCI). The principle the SCI is built on is, as we discussed in Chapter One, often associated with the principle that founded the success of the PageRank algorithm. Bernhard Rieder (2009, p. 137) describes Garfield's contribution in relation to information retrieval as a move "away from content-based organization towards topological analysis built on graph theory". Rieder (2009) continues with a comparison of how search engines moved from web search engines such as AltaVista, that built primarily on representation of content, to Google, that brought in the revolutionary principle of PageRank, where link analysis became the dominant organising principle. In this way, the work of Garfield is discussed as having laid the ground for the PageRank algorithm.

The information retrieval tradition has been described by White (2016, p. xii) as being structured by four "revolutions": the cognitive revolution, the relevance revolution, the interactive revolution, and finally the data revolution. Firstly, the cognitive revolution challenged so-called "system-oriented" information retrieval research in favour of "user-oriented" research. It is a view of users according to which their search behaviour primarily is dependent on their cognitive structures. That is, instead of focusing on document representation, the user-oriented information retrieval tradition focused on the user. Secondly, the relevance revolution is closely related to this cognitive understanding of information retrieval. Pia Borlund (2003, p. 923) describes this revolution: "The result of the relevance revolution is an increasing acceptance that relevance should be judged in relation to the information need rather than the request". When evaluating information retrieval systems, the relevance of information becomes a question for someone in a situation, rather than as a relation between index term and query. Thirdly, the interactive revolution revolved around an increasing understanding of so-called lookup searches as not being enough in order to understand information retrieval as carried out outside the laboratory. Instead, research focused more on series of searches in which relevance is conceptualised as a dynamic concept developed over time. Marcia Bates' (1989) *berrypicking model* plays a seminal role in visualising how users are seen to interact with the system during an information seeking process. Interactive information retrieval stresses how the trigger for information retrieval – what is labelled the information need – is dynamic and how it, just like relevance, develops and changes through interaction with the system. Finally, the data revolution, according to White (2016), is a result of the enormous amount of data available, not least due to the internet and the web. In fact, it is possible to say that the spread of general-purpose search engines for the web produced a new field of research – web search studies (e.g. Lewandowski 2012b). This research, which is developed in more detail below, builds to some extent on the information retrieval tradition, but the web also engenders new research questions. Not least, socio-economic questions about markets, business models, and monopoly supplementing

the more technically oriented information retrieval tradition become relevant (Lewandowski 2012a).

It is harder to identify a starting point for the information behaviour tradition. An interest in how people deal with information or information artefacts has attracted research from a number of different disciplines and goes back a long time. In information science, this tradition was originally often referred to as *user studies* and has as such been reviewed a number of times (Menzel 1966; Paisley 1968; Wilson 1994). The label information behaviour came about later and is defined by Tom Wilson (1999, p. 249) as "those activities a person may engage in when identifying his or her own needs for information, searching for such information in any way, and using or transferring that information". In this overview, we take a starting point for information behaviour research in the so-called user-centred tradition in information science and more specifically library and information science (LIS) from the late 1970s and early 1980s. At that time, a number of researchers broke with the dominant quantitative information retrieval tradition as well as with earlier user studies that dominated research in information science. They formulated an alternative research agenda that became known as INSU, short for *Information Needs, Seeking and Use*. However, it was not primarily a theoretical break from information retrieval, even if such early examples exist as well. The user-centred tradition started in the same theoretical climate that much profiled information retrieval research at the time was couched in – that is a cognitive perspective, or a cognitive viewpoint, which was dominant for many years. Nicolas Belkin's (1980) *Anomalous' State of Knowledge* (ASK) theory and Peter Ingwersen's *Cognitive viewpoint* (1992) were tone-setting and to some extent their influence can still be seen. The ASK model emphasises the experience of people when they lack information, which is called an "anomaly" in their state of knowledge. Such an anomaly creates a cognitive uncertainty that can be translated into an information need. In theory, the information retrieval system has the potential to provide a person with an anomalous' state of knowledge with a document that satisfies that person's information need.

According to the history of the user tradition, Brenda Dervin and Carol Kuhlthau are two early proponents taking as a starting point users and their perspective of the world rather than an information system. Brenda Dervin's (e.g. 1998) *Sense-Making Theory* starts from an interest in how people create meaning in problematic situations in which they experience a gap in their understanding. An information need is an expression of what information an individual needs, a bridge, to fill the experienced gap. Information seeking is thus something we as humans constantly carry on doing as a way to construct sense in our lives. Carol Kuhlthau (1991) has a similar starting point in the users' perspective, but with an empirical focus on information seeking in school settings. She developed a model of the information seeking process stretched over the time it takes to solve a school task. She identified a series consisting of six stages: initiation, selection, exploration, formulation, collection, and presentation. Kuhlthau attached feelings, thoughts and activities to each stage. The novelty of Kuhlthau's work lay with the way in which she

envisaged and emphasised the emotional part of information seeking, outlining how information seeking is far from just a purely intellectual and cognitive affair. This perspective helped the research community – as well as librarians – for instance to improve their understanding of pupils experiencing frustration when seeking information.

A popular point of reference in the user perspective tradition is the work of Robert S. Taylor (1968) who in his research on information seeking in libraries identified four levels of information need: 1. Visceral need, 2. Conscious need, 3. Formalised need, and 4. Compromised need. Taylor describes how an information need changes from something that cannot be expressed linguistically to, in the end, the point when the information need is translated to the required language of the information system at hand (classification codes, subject headings, thesaurus, etc.). For Taylor, what is important in information seeking happens inside the users. The most important lesson from Taylor's article is perhaps his way of describing the differences between an explicit or implicit information need where an implicit need is a need not yet recognised as such by the person supposed to be having it, the user of the system in other words. We can see the influence of Taylor's research in the relevance revolution in information retrieval that distinguishes between information needs and actual requests as expressed above by Borlund (2003). The step between formalised need and compromised need is of course of less importance when the use of a general-purpose web search engine is in focus in contrast to much of the analogue card information systems that existed in the 1960s, but the distinction reminds us how we in one way or another adapt to the system we use. Despite its popularity, Taylor's (1968) work on information needs, and its more implicit consequences for understanding relevance, has also been criticised (e.g. Hjørland 2010) for psychologising information needs. We expand on some of this critique in Chapter Three.

The so-called Information Needs, Seeking and Use (INSU) studies have incorporated a more contextual understanding of information throughout the years, with the bi-annual Information Seeking in Context Conference (ISIC) that started in 1996 as a clear indication of this development. To some extent, it is possible to say that the research in the information behaviour tradition was institutionalised through this very conference (Vakkari 2008). The then new research agenda not only broke with the traditional information retrieval research in methodological terms (from quantitative research to qualitative), it also broke with a narrow focus on particular information systems and their functions as such. There have been attempts to bridge the two traditions of information retrieval and information behaviour research in order to formulate one coherent research paradigm. Often, these attempts have been formulated by information retrieval researchers approaching them with a theoretical toolbox equipped with cognitive approaches. In Peter Ingwersen's and Kalervo Järvelin's *The Turn* (2005) and later in Ryen White's *Interactions with Search Systems* (2016), as well as in numerous other publications, information retrieval and information behaviour research are presented together, sometimes referred to as *Information Seeking and Retrieval* (IS&R). In a recurring

reference to Tom Wilson's (1999) "nested model", information behaviour, information seeking behaviour and information searching are concepts presented as three nested circles, with information searching behaviour in the inner circle, like a Russian *matryoshka* nesting doll. Yet, and attending to this more closely, it seems as if the major watershed in information retrieval and information behaviour research is not always the one distinguishing a near-sighted focus on interaction from a more far-sighted perspective providing a richer contextual understanding. Rather, the main difference seems to be between, on the one hand, research that has as its starting point in the premise of providing results that in a more instrumental way can be used for constructing better information systems and, on the other hand, research with a starting point in critical perspectives and people's practices, grounded in a less pragmatic premise. This difference relates to the slightly exaggerated, but still interesting question of whether information science research should have the instrumental purpose of developing better systems as its raison d'être, or whether it can and should also contribute with an analysis of other interactions with information than in information systems as well as with the analysis of cultural and political conditions these information systems derive from and shape.

Both information behaviour and information retrieval research are known to produce considerable numbers of models. Numerous models have been developed in order to describe or understand searching and predict future seeking and searching behaviours: from the simplest type of model illustrating how a document is represented in a database and matched by a query that is an outcome of an information need – to vastly complex models, which include myriads of different interrelated factors. David Bawden and Lyn Robinson (2012, pp. 192–197) group models in information behaviour research into four categories: descriptive models, process models, cognitive models, or complex models. Some of those aim at describing various factors involved, while others are closer to establishing testable relations between variables. They all strive to understand searching (or seeking) at a generalised level, even if they often include context, task, or situation as one aspect of many. Most often, models in the information seeking literature take an individual need (or subjective understanding of a task or a situation) as their starting point. David Ellis (1989) has presented a model with detailed features during the information seeking process: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, ending. Also, Kuhlthau's (1991) descriptive model of the information seeking process depicts a general process. Yet Kuhlthau's model, in addition to activities similar to those in Ellis' case, also involves feelings and thoughts. Another example, also mentioned above, of a more elaborate model of information searching is Marcia Bates' (1989) *berrypicking* model, which emphasises the interactive character of searching in information systems. Compared to a static model of information retrieval, which her model explicitly contrasts itself in relation to, Bates' model brings to the fore how what we are searching for, what we want to know, evolves during the process of searching and that peoples' information searching is a far less a systematic and logic behaviour than what much of the literature then seemed to suggest.

Relatedly, Nils Pharo (2004) proposed the "Search situation and transition model", which was developed specifically by studying web users during real-life web search sessions. It is a finely grained model that tries to account for complex work tasks and multi-facetted search situations. This attempt at systematising web search behaviour results in an information behaviour model which is formulated by accounting for concepts from information retrieval research (e.g. relevance) as well as the information seeking literature (e.g. work task). Through this approach a considerably richer understanding becomes possible of the interplay between the reason for a search and the varying ways in which results are judged and related to. That real-life web information searching is a complicated and sometimes messy matter becomes even clearer in the work Nils Pharo did together with Kalervo Järvelin (2006). They analysed textbooks on web information searching through the lens provided by empirical data of real-life web information search processes. The discrepancy between the normative and prescriptive ways of imagining web searching in the then available textbooks often captured in idealised models of rational behaviour and the data they present on what people really do, how they reason and also change what they do is striking.

Seeking for serious information

Regardless which model we look at, both the information behaviour and the information retrieval research tradition often have as their main focus information seeking for unknown documents or facts of "serious" character. Gary Marchionini (2006, p. 42) distinguishes between three types of searching activities: lookup, learning, and investigating. *Lookup* is in Marchionini's model exemplified with fact retrieval, known item search, navigation, transaction, verification, and question answering. *Lookup* searching is the type of information searching that information retrieval research has traditionally focused on with clearly defined needs that can be met with a document acquired through a correct search strategy. This includes questions that can be answered through one search and more complex questions that need a series of searches. In both cases, it is, however, a more or less defined problem that generates the activity of searching. *Learning* is, by Marchionini (ibid.) exemplified with knowledge acquisition, comprehension/ interpretation, comparison, aggregation/integration, and socialising. The third activity, *investigating*, is exemplified with accretion, analysis, exclusion/negation, synthesis, evaluation, discovery, planning/forecasting, and transformation. Lookup is the simplest form of activity, while Marchionini refers to learning and investigating as *explorative search*. He claims that, when people are confronted with an information demanding task, searching is increasingly carried out for exploration, which was not the case before the invention of the web. It is probably true that nowadays we take all kinds of questions to the web, whereas before we had other strategies involving different media and people.

That many of the various strategies for finding out about something and the use of different media have been more or less replaced by the use of web search

engines and social media, does not, however, mean to say that they have all merged into one and become identical. Differentiation happens elsewhere, not so much at the level of medium, but at the level of how the information infrastructure is given meaning and how it organises social practices. This claim will be developed over the course of the following chapters. Having said that, as we discuss in Chapter Four, it seems also as if, due to the logic of the web search engine, we transform more complex questions into simple ones when confronted with a search engine's specific affordances (Huvila 2013; Sundin et al. 2017). Marchionini (2006, p. 43) explains how the development from lookup to exploratory search has altered information retrieval research and advanced a growing interest in its interactive aspects:

> Rather than viewing the search problem as matching queries and documents for the purpose of ranking, interactive IR views the search problem from the vantage of an active human with information *needs*, information *skills*, powerful digital library *resources* situated in global and locally connected *communities* – all of which *evolve* over time.

Reijo Savolainen (2017) discusses the previously mentioned *berrypicking* model by Marcia Bates as an example of explorative search and thus provides an excellent example of how information behaviour research can contribute to the development of information retrieval and vice versa. He successfully accounts for the complexity of information searching where searching in order to solve a specified problem is only one type of search. Savolainen, probably due to his interest in everyday life, emphasises the importance of "curiosity" and "desire for learning" (2017, p. 2) as reasons for search in addition to the well-defined triggers that information retrieval studies often take as their starting point. As we will see in considerably more detail in Chapter Four, people's searching for information in everyday life is messy, ill defined, and simply not always suited to being expressed in detailed models, without losing much of its complexity.

What should be noticed in the examples provided by Marchionini and the information retrieval tradition is that they all refer to searching for documents that provide the searcher with some kind of epistemic content. Research where search has been associated with pleasure, amusement, or just passing time is absent in traditional information retrieval research, even if we above see how Savolainen's (2017) analysis of the *berrypicking* model is an attempt to fill explorative search with a content the information retrieval tradition normally does not do. At the same time, there is clear evidence, which we will get back to below, that the absolute majority of searches on the web concern precisely this – pleasure, amusement, or even just passing time (Waller 2011). Also, information behaviour research has had a focus on searching for information as epistemic content or representation of such content in documents. However, in the information behaviour tradition there are also exceptions to this general rule, as we will show further on.

Web search studies

The emergence of general-purpose web search engines has fostered new types of research questions, new data, new methods and new answers. Michael Zimmer (2010, p. 508) describes this fast-growing research field as the "meta-discipline of web search studies". To some extent, web search studies built on the results from information retrieval research, but there are also differences. In an introduction to web search studies, Dirk Lewandowski (2012a, p. 4), with reference to information retrieval research and library and information science, states:

> With web search engines, both communities are challenged, in that (1) other communities become more and more interested in search engine studies, (2) it becomes clear that only a deeper understanding of Web searching will suffice, which requires a combination of methods from different disciplines and (3) the social impact of Web search engines, which is only sometimes the focus of both disciplines, is an important area to consider.

Lewandowski (ibid.) argues for the need for an understanding of both web structures and search engines techniques as well as an understanding of users and their doings. This combined focus is also the starting point for this book. When search engines are accessible to everyone everywhere, new kinds of research questions become important. Traditionally, information retrieval has dealt with a homogenous type of documents in test databases, such as in the Text REtrieval Conference (TREC), but web search engines provide results including different formats, genres, and qualities. Web searches are often categorised in accordance with so-called query-intent, which is the (assumed) purpose of a search, the need it responds to. In an influential paper, Andrei Broder (2002; see also Jansen, Booth, & Spink 2008) categorised web search queries according to query intent as *informational*, *navigational*, or *transactional*. Informational web search refers to searches for a specific topic, where the searchers do not know what kind of document they need before accessing it. Navigational search then refers to when searchers know in advance which document or web site they are looking for, but use a search engine to access it instead of going there directly. For example, when you know you want to access a certain newspaper and you are using a search engine to find it and access it. Finally, transactional search describes all those searches whose goal is to buy or in other ways acquire products or services. Since it was first proposed this categorisation has become more fine-grained as new categories have been added in different studies (Lewandowski, Drechsler, & von Mach 2012). Yet as a basic structure these three categories are still influential.

Vivienne Waller (2011) analysed transaction logs to study types and topics of Google search queries. She found that in the material she analysed half of the searches submitted to Google concerned popular culture and ecommerce. Furthermore, half of the searches were of a kind "where it appeared that the searcher had no particular destination in mind" (Waller 2011, p. 770), as she formulates it.

Waller makes a case for not only seeing information searching as a way of solving knowledge-based problems, which is the way information retrieval in particular, but to some extent also information behaviour research, has traditionally been framing search. Instead, Waller argues for the leisure-dimension of searching: "the Internet search engine is not only an interface to information or a shortcut to Websites, it is equally a site of leisure" (Waller 2011, p. 774). On the basis of Waller's research, it seems safe to say that the dominance of interest in pre-web information retrieval – the focus on search of epistemic content or documents – does not always translate so well to web search. Waller's findings have led Ralph Schroeder (2015, 2018) to conclude that too much research interest has gone into evaluation of web search results based on the tradition of information retrieval. According to Schroeder, leisure sites cannot be evaluated in the same way as, for example, information on health or other issues, which are seen to be of higher social importance. This is a valid point, but at the same time, even if the number of searches on popular culture or ecommerce by far outnumbers searches on "socially important" content (Schroeder 2015, p. 157), they still constitute, as Waller (2016, p. 792) reminds us, "an incredibly large *number* of searches, given that Google receives 40,000 search queries every second". In fact, as an aside and to add further perspective, in the two years after this claim was made, this already unfathomably large number has increased by 50%, and in June 2018 Google alone was estimated to have received over 3.7 million searches on average each minute (Domo n.d.).

Another line of research investigates what people choose from the search engine results page – often referred to as SERP – and why they choose as they do. This work convincingly shows that how people choose links is primarily based on where these links are located on the search engine results page. White (2016, p. 65) refers to this as position bias, "a case in which a searcher's decision on which result to select is based on rank position of the search result rather than its relevance to the information need. This is also referred to as 'trust' bias or 'presentation bias'". For example, Bing Pan and his colleagues (2007) in an eye tracking study show how people tend to click on the first links of the search engine results page. They conclude: "when all factors are considered, subjects trust Google's positioning more than their rational judgements based on evaluation of different alternatives" (ibid., p. 816). Ten years on Sebastian Schultheiß, Sebastian Sünkler, and Dirk Lewandowski (2018) replicated Pan et al.'s study with some modifications. They confirmed that the order of results determines how people look at results pages. Yet relevance is a decisive criterion for which links they choose. However, this has to be put in relation to another early finding, namely not only do most people not go beyond the first results page presented to them, they never even consider all results on the first page. Nadine Höchstötter and Dirk Lewandowski (2009) demonstrate empirically that people most often do not scroll down below "the fold", that is what the user sees on the screen without having to scroll. In other words, the call for users to assess the quality of information when searching the web is in reality often left unheard. There seems to be a gap between the recommendations by teachers, librarians, and other professionals who instruct people about how to

search and how people actually go about searching. This gap is discussed below, and further explored in Chapter Five, but for now it is enough to point to the fact that web search engines deliver a huge number of results of divergent character, and users most often trust the work of the search engine and therefore often chose links from the top of the search engine results page.

The diversity of quality is not really considered in commercial web search engines. Most of us when searching primarily are interested in a combination of topicality and a link high up on the search engine result page. Media scholars Ken Hillis, Michael Petit and Kylie Jarrett (2013, pp. 58–60) show how Google relates quality and relevance and how it is doing its best to furnish an understanding of search results as objective. Is another understanding of search engine results possible? Relatedly, Yvonne Kammerer and Peter Gerjets (2012) argue how a different search engine result page than the one based solely on ranking of relevance could provide better affordances for people when choosing links. The authors posit that a search engine results page that visualises the diversity of types, ideas, genres, formats, and so on of websites with the same topic provides users with cues to make more informed or at least differently informed link choices. Kammerer and Gerjets (2014) tested a grid interface of presenting the results from search and their study confirm the hypothesis that a different structure of interface leads the users in different directions, away from the tendency to select only links at the top.

Let us now alter the perspective, from a specific use of search engines, to information practices and seeking in everyday life.

Information practices and information seeking in everyday life

Not much empirical attention has been paid to the empirical investigation of the use of search engines in everyday life, at least not in information science. One reason could be the division of labour that has been established in information science. Questions pertaining to searching by means of search engines have so far mostly been investigated by researchers coming from backgrounds in which quantitative methods dominate. At the same time, researchers with an interest in everyday life and its complexity and richness, with a predilection for qualitative methods and a starting point in practices, have rarely demonstrated an interest in search engines.

A couple of early exceptions, where search engines actually have been investigated in everyday life, from the early years of this century are the work of Anders Hektor (2001) and Soo Young Rieh (2004). These two studies took off when the home computer was connected with a modem that made searching domesticated for the first time. Hektor (2001) investigated peoples' information behaviour in everyday life using interviews and diaries. He identifies nine different types of information activities, where "search and retrieve" is one of them. Despite the focus on everyday life, the starting point in Hektor's study is a problem that generates a query. The informants in his study searched either for reference information or market information. Rieh (2004) explored web searching qualitatively in

another early attempt to explore what actually happened when online search facilities were made available in people's everyday lives. She asked her informants to fill in a structured search diary that collected what they searched for in detail, how they started their search, and whether they were successful or not. The short search diaries were followed up by an interview that, among other things, tried to understand the informants' goals when searching the web. Most goals were concrete and specified, but Rieh (2004, p. 751) also found that "subjects did not always initiate the search process because they had specific information problems to be solved". This statement constituted in fact a break with a long-standing information science tradition that most often starts from a task or a problem that generates a definable information need (Talja & Nyce 2015). Since these early studies the information infrastructure has, as we know, changed considerably. With information access having become mobile, the smartphone accompanying us everywhere and smart speakers having moved into homes, searching is not just done at work or at home, but also in-between and potentially almost everywhere else.

Reijo Savolainen has in a number of studies explored information seeking – not just in relation to online search – in everyday life and already in 1995 he coined ELIS (Everyday Life Information Seeking) as a way of making visible the information seeking outside of work or research. At the time, work focusing on researchers' or professionals' seeking and searching by far outnumbered studies of information seeking outside work or school. Ever since then, however, a stream of research has investigated information seeking as it takes place in the daily routines of people (e.g. Erdelez 1997; Kari & Hartel 2007; Williamson 1998). Pamela McKenzie (2003, pp. 19–20) argues that the many models of information seeking do not provide an understanding of information seeking in everyday life. She highlights three reasons for this shortcoming; "current models tend to focus on active information seeking, to the neglect of less-directed practices", "many research-based models of information seeking are derived from studies of scholars or professionals" and "many models have been developed using a cognitive approach to model building". The cognitive model, according to McKenzie, only furnishes our grasp of the role of searching in everyday life to a limited extent. McKenzie tries instead to anchor her research in the complexity of social practices.

The so-called practice turn in social and cultural sciences has also clearly affected information science. Pamela McKenzie (2003), Kimmo Tuominen, Reijo Savolainen and Sanna Talja (2005), Tiffany Veinot (2007), Annemaree Lloyd (2010, 2014), Andrew Cox (2012), Johanna Rivano Eckerdal (2012), Ola Pilerot (2013) and many others have built a rich theoretical and empirical understanding of the role of information in various practices and in the everyday lives of people, either in or outside work. This stream of research has brought light to how information forms part of all social practices, rather than focusing on certain information practices, as distinct from other practices. Cox (2012, p. 185) formulates this in the following way:

> [I]nformation activities are woven through all social practices, and this is even more evident today through ubiquitous access to information resources through the internet. Thus we need to look at the information aspect of all

> social practices. Escaping a narrow preoccupation with goal-oriented information seeking, we need to first ask within any practice what, for social actors, constitutes information, and then how do they find, use, create and share it.

The practice turn has brought new oxygen into information behaviour research. Information has been studied as emergent from specific ways of doing things and entangled across various objects, most significantly the human body (Lindh 2015, 2018; Lloyd 2010, 2014; Rivano Eckerdal 2012), but also for example compost heaps (Haider 2011) or vault inspection forms (Veinot 2007). Not least Lloyd's work on supplementing coded knowledge in information literacy research with sensitiveness also to corporeal knowledge, with inspiration primarily from Theodore Schatzki, has been influential. Talja and Nyce (2015) skilfully develop how situations and tasks can be framed within a practice theoretical understanding and thus contribute to a rich understanding of the roles of information and information related activities in everyday life. Clearly, as their work shows, practice theory in information science cannot be reduced to one approach but its incarnations represent slightly different understandings of social practices, yet with a similar basis (see also Pilerot et al. 2017).

In the practice approach, information is everywhere, but at the same time not much attention has been paid to where information science has traditionally been looking – in information systems. When taking a starting point in peoples' practices rather than in a search engine or other resources, the search engine takes a back seat. Practice theory has greatly contributed to information science. Yet, it has concomitantly made one of information science's established areas of expertise – use, evaluation and development of information retrieval systems, now mostly search engines – less visible in the broader debate, at a time when general-purpose web search engines and similar search tools are no less than changing our culture and society. Paradoxically, the very research in information science that leans on theories from the social sciences and humanities has more or less dropped the field's traditional interest in search and search engines at the same time as researchers from other social sciences and humanities disciplines have started to develop a keen interest in search and search engines. Exceptions of course exist. In addition to the above-mentioned work of Rieh (2004) and that of Hektor (2001), more recently Cecilia Andersson (2017a, 2017b) has shown how Google is understood differently depending on the context within which it is used (Andersson 2017b). In school, Google is primarily referred to as a tool for finding facts and outside of school, even though it is used, it remains largely invisible. In both contexts however, Google is accredited with a lot of trust (Andersson 2017b). In Chapters Four and Five, we address this paradox.

Jack Andersen (2018) makes a theoretical claim for how search engines and their use are increasingly submerged into everyday life. Making communications recognisable to algorithms, he argues, makes them "part of our habitualized actions" (Andersen 2018, p. 1144). That is, people orient themselves towards the algorithms underlying search or at least to their understanding of them and of how they work.

The extent to which search and search engines are embedded and constitutive of everyday life is thus hard to understate. "Search engines", writes Andersen "have given shape to search as a habitualized form of communicative action not present before". And he continues:

> In one single medium, they offer speedy access to entertainment, goods, news, and information on a scale and to an extent not seen before in the history of media and communication. In addition, today search engines or a search function is part of other forms of social, networked, or mobile media, thereby institutionalizing search.
>
> *Andersen 2018, p. 1141*

In the next section, we discuss in more depth how search and search engines have been investigated as social and cultural artefacts, which are at the same time shaping and being shaped by social, cultural and material conditions.

Perspectives on the social shaping and the politics of search engines

Hillis, Petit and Jarrett (2013) elucidate in their book *Google and the Culture of Search* how Google has come to shape not just the way we find information, but our culture: "Google has achieved what we argue is its socially *consecrated* status" (p. 7). With such a status it is no wonder the search results are taken for granted. Yet how are the search results shaped and how are search engines embedded in society? As stated above, since the spread of web search, a research interest in search engines and particularly in Google has developed outside information science. This research fills a gap, since information science so far often has not so much considered cultural or social and certainly not socioeconomic or political aspects when investigating online search (for exceptions, see e.g. Andersson 2017a, 2017b; Andersen 2018; Huvila 2016; Noble 2018; Sundin et al. 2017; Zimmer 2008). In their overview of research on information searching and retrieval, Bernard Jansen and Soo Young Rieh (2010, p. 1530) note:

> while information searching researchers have moved into areas of organizational, cultural, and social contexts affecting or moderating the benefits of information use, information retrieval researchers have paid little attention to social aspects of information use.

In the following, we present selected research on social shaping of search engines and the politics of search, as well as on search as politics. The issue that structures the majority of the literature we discuss here evolves around the question of search engine bias, either explicitly and measurably or as an issue to engage with conceptually. The research is grounded in different disciplinary traditions, including information science.

In 2000 Lucas Introna and Helen Nissenbaum published their greatly influential article, "Shaping the web: Why the politics of search engines matters". The authors relate search engines to politics while making two claims: 1. Search engines are biased in the way they include and exclude sites as well as in the way they rank these sites, and 2. Search engines are too important for society. They cannot be left solely to the market. Instead, search engines ought to be regarded as a public good and algorithms should be made transparent. Introna and Nissenbaum conclude their work by stating that "search-engine design is not only a technical matter but also a political one" (2000, p. 181). Since then, a considerable number of researchers have followed up on this motif and have tried to understand the social shaping of search engines, the politics of search and search engine bias (e.g. Dutton et al. 2017; van Dijck 2013; Gillespie 2017; Goldman 2008; Halavais 2018[2009]; Hillis, Petit, & Jarrett 2013; Hinman 2008; Jiang 2014; Mager 2012; Noble 2018) and also provided meta-perspectives on and summaries of the existing research (Granka 2010; Schroeder 2018). For the sake of brevity, we focus on a small selection of relevant, mostly empirical work to establish in broad strokes how search engine bias and similar issues play out.

In a study of differences between Google's server in Hong Kong and the Chinese search engine Baidu, Min Jiang (2014) visualises how the two search engines differ in accessibility, overlap, ranking, and bias. The author found that there is only an overlap of 6.8 %. She also found possible bias patterns, for example in relation to how the search engines treat Wikipedia and the Chinese collaborative online encyclopaedia Baidu Baike. Where Google seems to favour Wikipedia, Baidu seems to favour Baidu Baike. Jiang maintains in a concluding remark, "these results suggest that search engines can be architecturally altered to serve political regimes, arbitrary in rendering social realities and biased toward self-interest" (ibid., p. 229). Thus, in the case of the comparison of Google and Baidu undertaken by Jiang (2014), the results show a very small overlap. That means that a searcher using one search engine gets an almost completely different understanding of what there is to know compared to a searcher using the other search engine if they use them in the same way. In comparison, Andrei Zavadski and Florian Toepfl (2019) study web searching as a mnemonic practice and compared how Google and the Russian search engine Yandex represent Russian historical events. They show how both search engines "reproduce and reinforce the dominant narratives supported by the ruling elites" (ibid., p. 21). Russian pro-regime media was prioritised and independent media were less present both in Google and Yandex. We might however ask whether, for example, Google's schematic prioritising of authoritative institutions can lead to entirely different results depending on the form of government? The majority of the so-called alternative media are in most contemporary liberal democracies often found on the far right, while in an authoritarian state, the alternative media might be what is considered the mainstream media in other countries (see also Lev-On 2008). That is, bias is a question of perspective, and different types of biases are at work at the same time, but these perspectives are hard to see in a search engine results page. Search engines never declare their

standpoint; it is assumed and maybe users assume they are the ones providing it. Another aspect of search and search engines in relation to global geographical differences is who produces the content and how the content is represented in searches. Andrea Ballatore, Mark Graham, and Shilad Sen (2017, p. 1211) show how "wealthy and well-connected countries tend to have much more locally produced content that is visible about them than poor and poorly connected countries". The authors relate to this fact as "digital hegemony" where particularly countries in the Global South are defined by the Global North.

We want to approach the argument concerning search engine bias from a slightly different angle. In a way, the entire discussion of whether search engines are biased is misleading, since the answer can never be no. There is no possible way to provide a *neutral* or *objective* search engine. The question has to be how are they biased and how visible are the various biases (in the plural)? The answers necessarily have to be complex and take into account that search engines constantly change, that they are but one part in a complex system of various power relations, and that bias is not a static part of the system, but happens also at the point of use. In fact, the presentation of Google as "an objective courier of online information" needs to be seen as little more than a successful marketing strategy (Bilić 2016, p. 7), not something to strive for. If there is no neutral point of objectivity, a stable ground of representation, the discussion of whether or not there is bias is distracting. Search engines always provide us with a version of reality. Thus, what is interesting and relevant is to map more precisely which or whose values are represented and how.

In her book *Algorithms of Oppression*, Safiya Umoja Noble (2018) enters a conversation with a broad array of research – empirical and conceptual – on the issue of search engine bias, formulated in a variety of ways, and work engaging critically with various hegemonic and cultural assumptions built into information systems. Based on strong empirical evidence, Noble elucidates how racism and sexism contribute to organising the results of Google both due to the way algorithms are encoded and due to what other people search for and how society is organised. She bases her work on a number of convincing cases, for example a comparison between the results from searching for "black girls" compared to "white girls", where the query "black girls" turned up sexualised photos in a way that the query "white girls" did not. In this case, the presumed "democracy" of the search engine is at least one cause of the problems, at least if democracy is understood as majority rule. If most people's search behaviour and the world they search within is racist and sexist, the search engine will favour racism and sexism and by doing so making it more and more visible, thus further reinforcing it. In addition, when data making up the infrastructure for searching, providing, for example, users with racist and sexist autocomplete suggestions, is prejudiced against certain groups it will reinforce and strengthen the same bias. Yet, as we will develop in more detail below, users of search engines are implicated in this loop.

Obviously, democracy is in practice not simply majority rule nor are search engines mere representations of what is most popular, what gets the most clicks or links

(see also Diaz 2008; Rieder 2009). Rather, at least in liberal democracies, in order to ensure representation, some kind of protection of minority rights is institutionalised in different mechanisms, not least to guard democracy from and against its own demise through strict majority rule and also with a view to a future to strive for. Search engines have mechanisms, including automated but also human content moderation and rating, that is carried out according to detailed policies, in order to filter what is allowed to surface and what not (see also Noble 2018, pp. 56–58 and on the work of human raters, see Bilić 2016). Engin Bozdag (2013) shows how the gatekeeping function, traditionally attributed to human beings, in contemporary information services has been supplemented by algorithmically framed gatekeeping. However, according to Bozdag, algorithmic gatekeeping should be seen as being made up from a combination of humans and machines. To be exact, he makes visible how, for example, Google actively effects the ranking by manually downgrading a link to a website. If the term *biased* is to be used, then this manual editorial work also needs to be framed as a way of tweaking bias in the search results, but not as one of making it more neutral or objective. Thus, as Noble (2018, p. 148) points out, what is at stake when we discuss the issue of search engine bias and the relationship between search engines and information presented through their use is the fact that "[s]earch does not merely present pages but structures knowledge, and the results retrieved in a commercial search engine create their own particular material reality". Specifically, she continues, "[r]anking is itself information that also reflects the political, social, and cultural values of the society that search engine companies operate within, a notion that is often obscured in traditional information science studies".

In a series of articles, Astrid Mager (2009, 2012, 2017; Eklöf & Mager 2013) demonstrates how search engines co-construct knowledge and how they are not just stabilised by themselves, but are entangled with the practices of content providers and users. A similar point is made by Tarleton Gillespie (2017) when he stresses the active role of information producers in order to become visible in search engines. Mager (2009) also maintains that Google has become an "obligatory passage point" that both web providers and users need to adjust to (see also Sundin & Carlsson 2016). By drawing on van Couvering (2008), Mager (2012) claims that search engines have gone from an academic product, with an origin in the information retrieval tradition, to a commercial product where providing search has become a way of selling advertisements rather than the other way around. As one of her interviewees states, "Google is not just search, in fact Google is not primarily search, it's advertising, right?" (ibid., p. 776). Google personalises search results to adapt them for individual users, even though the extent of the personalisation does not yet seem to be overwhelming (e.g. Hannak et al. 2013), but it is the personalisation of advertisements that Google makes possible that has bigger consequences. Mager identifies how search engines are entangled with a capitalist ideology in which different actors compete over visibility. The lack of transparency makes it almost impossible for citizens to understand how, for example, scientific controversies are played out by Google Search, such as in the case of

biofuel as discussed by Eklöf and Mager (2013). The way Google, through its relevance calculation, co-produces what knowledge is expected to be is an oft-repeated argument in the literature (Hinman 2008; Rogers 2013; van Dijck 2010). The order of knowledge presented by the search engines legitimises the importance of certain knowledge claims.

While the issue of search engine bias is a much discussed and researched area, the relationship between search engine bias and personal bias is not. Zavadski and Toepfl (2019, p. 33) highlight the significance of the search terms used for which version of history surfaces. "This is most obvious" they write illustratively, "in the outcomes of the queries 'Annexation of Crimea' and 'Incorporation of Crimea', which – despite referring to the same historical event – produced fundamentally different results that tended to transfer recollecting individuals to webpages that actualised opposing memories of the event". Relatedly, the issue of political bias and Google and their relationship gets a very different twist in a study by Francesca Tripodi (2018) on how conservative Christians in the USA go about looking for information and how they interpret it, and mostly for politically relevant information. While there is a general consensus in the group on the anti-conservative bias of the mainstream media, this accusation does not extend to Google. In Tripodi's study, Google appears as a neutral source of information with people and organisations literally "putting their faith in Google" (Tripodi 2018, p. 32). However, Tripodi also notices how for one thing, this perception of the search engine as a neutral broker of facts and information is to a large degree based on a misunderstanding of how the search engine ranks results and the influence the searchers' own actions have on the results displayed. Specifically, she shows how "*the phrase someone Googles dramatically affects the information they receive*" (Tripodi 2018, pp. 29–30). In other words, Google is used to confirm someone's viewpoint by unreflectingly using phrases that originate in the ideological belief system a person is grounded in. When no alternative viewpoints turn up in the search results, this can be taken as a confirmation that they do not exist, since Google is positioned as neutral, even if this neutrality is just assumed and never actively postulated. The significance of how search queries are formulated can also be actively exploited, not least by extremists, who identify what Michael Golebiewski and dana boyd (2018, p. 1) call "data voids". These are topical areas about which very little or even no information is published online. Once identified – often in response to an emergency event or in relation to problematic (typically discriminatory and prejudiced) search terms – these voids can be filled with whichever content someone wants to push. Users are then directed to these newly filled voids by mentioning phrases or keywords to be googled in social media or suitable forums. Since no other information exists that corresponds as well to the query as this newly created information, it will appear very high up on the results page and seem relevant and in all likelihood trustworthy. Golebiewski and boyd (ibid., p. 3) mention the example of white supremacist killer Dylann Roof, who in 2015 shot dead nine African Americans in a church in the USA. In his "manifesto" he described how he had read an entry on Wikipedia which, in his words,

"prompted me to type in the words 'black on white crime' into Google, and I have never been the same since that day". In this case, a data void for a phrase common only in white supremacist circles, had been filled with extremist content, thus pushing an extremely racist agenda via Wikipedia and Google into the informational texture of mainstream society. If we consider this together with the notion of a digital hegemony advanced by Ballatore, Graham, and Sen (2017) then it becomes clear that although the amount of data that can potentially be searched is unfathomably large it is not so for all possible queries. Thus, the fact that content can be and also is produced in direct response to how search engines work has very direct implications for the shaping of information, which in turn makes it obvious just how inseparably entangled online search and online information actually are and how significant the role of search engines is for *factualising* information.

Search engines act as gatekeepers between users and what there is to know. The search engine is a different type of gatekeeper from traditional intermediaries, for example in the form of librarians. Niels Kerssens (2017) investigates how search engines developed ideologically from the mid-1970s to the late 1980s. He insightfully sketches how the ideology of search during this period changed from a humanistic ideology, based in librarianship and library science, to a positivistic ideology, based in computer (and information) science. In the last two decades, library and information science has begun to develop a sensitivity to power in relation to material selection and classification (see, for example, Olson 2001), something which Noble (2018) also notices. However, the change that Kerssens describes from a humanistic to a positivistic ideology underpinning search, also brought about a change in where power relations are located and can be caught sight of. More specifically, as "editorial control" became increasingly hidden for the user, the user lost control: "the shift from human towards friendly software intermediaries increased the control of the computer system over the search process by disconnecting information retrieval in the back-end (a diminished form of) *selection powers* of users at the front-end" (Kerssens 2017, p. 229). A similar point is made by Rieder (2009), who argues that the principle of convenience competes with the one of autonomy. He makes the claim that, when the mediation process in searching is hidden, the autonomy of the searcher is getting lost. These conclusions are important for the understanding of the ideology shaping contemporary web search engines as tools that effectively provide their users with very little control. These considerations are especially relevant to have in mind regarding the increasing absorption of search engines into the ordinary things of everyday life and into the voices of AI assistants.

Chapter conclusion: Search as a social, cultural, and material/technological phenomenon

This chapter provides the basis for understanding search as a social, cultural, and material/technological phenomenon, which is expanded on in the following chapters. We provide a background with literature from information science as

well as from neighbouring disciplines. Information science is a discipline with fuzzy borders that is located in the humanities as well as in social science or even computer science. When situated closer to the humanities, it is often referred to as information studies or linked to libraries as in the compound library and information science. This fuzziness can, when investigating such a complex phenomenon as search, be regarded as a strength – search needs to be studied from different viewpoints. When search and search engines are investigated within information science it is often search and search engines *in general* that are the focus. The background from information retrieval gives the discipline a long-standing tradition in investigating search and search engines in different contexts. When information searching and search engines are investigated from media studies, communication studies, organisational studies, or STS, the focus is often different. These and similar disciplines do not have the same tradition and their researchers often approach search and search engines largely as a Google phenomenon. References to Baidu, Yandex and even alternative web search engines such as DuckDuckGo are sometimes made, but almost always with Google as a reference point. For us, it is obvious how the different ways of approaching online search benefit from each other. Online search is a social, cultural, and material/technical phenomenon. Within information science, we have presented information retrieval and information behaviour as different, although to some extent overlapping fields of research. The link between information retrieval and information behaviour research has often been the cognitive or psychological interest in searching. This research has provided detailed investigations of people's experiences of relevance as well as discussions of information needs. The tradition of information practices with its starting point in social science-oriented practice theory and its interest in everyday life has the potential to bridge information science research and other disciplines' investigations on search and search engines, but, as we have argued, this tradition has not so far given search and search engine use much interest.

References

Andersen, J. (2018). "Archiving, ordering, and searching: Search engines, algorithms, databases, and deep mediatisation." *Media, Culture & Society*, 40(8), 1135–1150.

Andersson, C. (2017a). "The front and backstage: Pupils' information activities in secondary school." *Information Research*, 22(1), CoLIS paper 1604. Retrieved from http://InformationR.net/ir/22-1/colis/colis1604.html.

Andersson, C. (2017b). "'Google is not fun': An investigation of how Swedish teenagers frame online searching." *Journal of Documentation*, 73(6), 1244–1260.

Ballatore, A., Graham, M., & Sen, S. (2017). "Digital hegemonies: The localness of search engine results." *Annals of the American Association of Geographers*, 107(5), 1194–1215.

Bates, M. J. (1989). "The design of browsing and berrypicking techniques for the online search interface." *Online Review*, 13(5), 407–424.

Bawden, D. and Robinson, L. (2012). *Introduction to information science*. London: Facet Publishing.

Belkin, N. J . (1980). "Anomalous states of knowledge as a basis for information retrieval." *Canadian Journal of Information Science*, 5, 133–143.

Bilić, P. (2016). "Search algorithms, hidden labour and information control." *Big Data & Society*, 3(1), 1–9.

Borlund, P. (2003). "The concept of relevance in IR." *Journal of the American Society for Information Science and Technology*, 54(10), 913–925.

Bozdag, E. (2013). "Bias in algorithmic filtering and personalization." *Ethics and Information Technology*, 15(3), 209–227.

Broder, A. (2002). "A taxonomy of web search." *ACM SIGIR Forum*, 36(2), 3–10.

Case, D. & Given, L. (2016). *Looking for information: A survey of research on information seeking, needs, and behavior* (4th ed.). Bingley: Emerald.

Cox, A. M. (2012). "An exploration of the practice approach and its place in information science." *Journal of Information Science*, 38(2), 176–188.

Darnton, R. (2009). "The library in the information age: 6000 years of script." In K. Becker & F. Stadler (Eds), *Deep search: The politics of search and beyond Google* (pp. 32–44). Innsbruck: Studien Verlag.

Dervin, B. (1998). "Sense-making theory and practice: An overview of user interests in knowledge seeking and use." *Journal of Knowledge Management*, 2(2), 36–46.

Diaz, A. (2008). "Through the Google goggles: Sociopolitical bias in search engine design." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 11–34). Berlin, Heidelberg: Springer.

Domo (n.d.). "Media usage in an internet minute as of June 2018." Statista – The Statistics Portal. Retrieved from https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute.

Dutton, W. H., Reisdorf, B., Dubois, E. & Blank, G. (2017). "Search and politics: The uses and impacts of search in Britain, France, Germany, Italy, Poland, Spain, and the United States." Quello Center Working Paper No. 5–1-17. Retrieved from https://ssrn.com/abstract=2960697.

Eklöf, J. & Mager, A. (2013). "Technoscientific promotion and biofuel policy: How the press and search engines stage the biofuel controversy." *Media, Culture & Society*, 35(4), 454–471.

Ellis, D. (1989). "A behavioural model for information retrieval system design." *Journal of Documentation*, 45(3), 171–212.

Erdelez, S. (1997). "Information encountering: A conceptual framework for accidental information discovery." In P. Vakkari, R. Savolainen, & B. Dervin (Eds), *ISIC '96 Proceedings of an International Conference on Information Seeking in Context, Tampere, Finland* (pp. 412–421). London, UK: Taylor Graham Publishing.

Frohmann, N. (2008). "The role of facts in Paul Otlet's modernist project of documentation." In W. B. Raward (Ed.), *European modernism and the information society: Informing the present, understanding the past* (pp. 75–88). Aldershot: Ashgate.

Gillespie, T. L. (2017). "Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem." *Information, Communication & Society*, 20(1), 63–80.

Goldman, E. (2008). "Search engine bias and the demise of search engine utopianism." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 121–133). Berlin, Heidelberg: Springer.

Golebiewski, M. & boyd, d. (2018). "Data voids: Where missing data can easily be exploited." *Data & Society*. Retrieved from https://datasociety.net/output/data-voids-where-missing-data-can-easily-be-exploited/.

Granka, L. A. (2010). "The politics of search: A decade retrospective." *The Information Society*, 26(5), 364–374.

Haider, J. (2011). "The environment on holidays or how a recycling bin informs us on the environment." *Journal of Documentation*, 67(5), 823–839.

Halavais, A. (2018 [2009]). *Search engine society* (2nd ed.). Cambridge, UK: Polity Press.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). "Measuring personalization of web search." In *Proceedings of the 22nd international conference on World Wide Web* (pp. 527–538). ACM.

Hektor, A. (2001). *What's the use? Internet and information behavior in everyday life*. Linköping, Tema: Linköping University.

Hillis, K., Petit, M., & Jarrett, K. (2013). *Google and the culture of search*. New York: Routledge.

Hinman, L. M. (2008). "Searching ethics: The role of search engines in the construction and distribution of knowledge." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 67–76). Berlin, Heidelberg: Springer.

Hjørland, B. (2010). "The foundation of the concept of relevance." *Journal of the American Society for Information Science and Technology*, 61(2), 217–237.

Höchstötter, N.& Lewandowski, D. (2009). "What users see: Structures in search engine results pages." *Information Sciences*, 179(12), 1796–1812.

Huvila, I. (2013). "In web search we trust? Articulation of the cognitive authorities of web searching." *Information Research*, 18(1) paper 567. Retrieved from http://InformationR.net/ir/18-1/paper567.html.

Huvila, I. (2016). "Affective capitalism of knowing and the society of search engine." *Aslib Journal of Information Management*, 68(5), 566–588.

Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.

Ingwersen, P. & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dordrecht: Springer Verlag.

Introna, L. D., & Nissenbaum, H. (2000). "Shaping the web: Why the politics of search engines matters." *The information society*, 16(3), 169–185.

Jiang, M. (2014). "The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China." *New Media & Society*, 16(2), 212–233.

Jansen, B. J. & Rieh, S. Y. (2010). "The seventeen theoretical constructs of information searching and information retrieval." *Journal of the American Society for Information Science and Technology*, 61(8), 1517–1534.

Jansen, B. J., Booth, D. L., & Spink, A. (2008). "Determining the informational, navigational, and transactional intent of web queries." *Information Processing & Management*, 44(3), 1251–1266.

Juel Vang, K. (2013). "Ethics of Google's Knowledge Graph: some considerations." *Journal of Information, Communication and Ethics in Society*, 11(4), 245–260.

Kammerer, Y. & Gerjets, P. (2012). "Effects of search interface and Internet-specific epistemic beliefs on source evaluations during web search for medical information: An eye-tracking study." *Behaviour & Information Technology*, 31(1), 83–97.

Kammerer, Y. & Gerjets, P. (2014). "The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface." *International Journal of Human–Computer Interaction*, 30(3), 177–191.

Kari, J. & Hartel, J. (2007). "Information and higher things in life: Addressing the pleasurable and the profound in information science." *Journal of the American Society for Information Science and Technology*, 58(8), 1131–1147.

Kerssens, N. (2017). "When search engines stopped being human: Menu interfaces and the rise of the ideological nature of algorithmic search." *Internet Histories*, 1(3), 219–237.

Kuhlthau, C. C. (1991). "Inside the search process: Information seeking from the user's perspective." *Journal of the American Society for Information*, 42(5), 361–371.

Lev-On, A. (2008). "The democratizing effects of search engine use: On chance exposures and organizational hubs." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 135–149). Berlin, Heidelberg: Springer.

Lewandowski, D. (2012a). "New perspectives on web search engine research." In D. Lewandowski (Ed.), *Web search engine research* (pp. 1–16). Bingley, U.K.: Emerald.

Lewandowski, D. (Ed.) (2012b). *Web search engine research*. Bingley, U.K.: Emerald.

Lewandowski, D., Drechsler, J., & Von Mach, S. (2012). "Deriving query intents from web search engine queries." *Journal of the American Society for Information Science and Technology*, 63(9), 1773–1788.

Lindh, K. (2015). *Breathing life into a standard: The configuration of resuscitation in practices of informing*. Lund: Lund Studies in Arts and Cultural Sciences.

Lindh, K. (2018). "The Construction of Lay Rescuers in Bystander CPR Classes." *Library Trends*, 66(3), 315–328.

Lloyd, A. (2010). "Corporeality and practice theory: Exploring emerging research agendas for information literacy." *Information Research*, 15(3) CoLIS794. Retrieved from http://InformationR.net/ir/15-3/colis7/colis704.html.

Lloyd, A. (2014). "Following the red thread of information in information literacy research: Recovering local knowledge through interview to the double." *Library & Information Science Research*, 36(2), 99–105.

Lovink, G. (2009). "Society of the query: The Googlization of our lives." In K. Becker & F. Stadler (Eds), *Deep search: The politics of search and beyond Google* (pp. 45–53). Innsbruck: Studien Verlag.

McKenzie, P. J. (2003). "A model of information practices in accounts of everyday-life information seeking." *Journal of Documentation*, 59(1), 19–40.

Mager, A. (2012). "Algorithmic ideology: How capitalist society shapes search engines." *Information, Communication & Society*, 15(5), 769–787.

Mager, A. (2009). "Mediated health: sociotechnical practices of providing and using online health information." *New Media & Society*, 11(7), 1123–1142.

Mager, A. (2017). "Search engine imaginary: Visions and values in the co-production of search technology and Europe." *Social Studies of Science*, 47(2), 240–262.

Marchionini, G. (2006). "Exploratory search: From finding to understanding." *Communications of ACM*, 28(9), 41–46.

Menzel, H. (1966). "Information needs and uses in science and technology." In C. Quadra (Ed.), *Annual Review of Information Science and Technology* (pp. 41–68), Volume 1. Wiley: New York.

Nicholas, D., Rowlands, I., Clark, D., & Williams, P. (2011). "Google generation II: Web behaviour experiments with the BBC." *Aslib Proceedings*, 63(1), 28–45.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Olson, H. A. (2001). "The power to name: Representation in library catalogs." *Signs: Journal of Women in Culture and Society*, 26(3), 639–668.

Paisley, W. (1968). "Information needs and uses." In C. A. Cuadra (Ed.), *Annual Review of Information Science and Technology (ARIST)*, Volume 3 (pp. 1–30). Chicago: Encyclopedia Britannica.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). "In Google we trust: Users' decisions on rank, position, and relevance." *Journal of Computer-Mediated Communication*, 12(3), 801–823.

Pharo, N. (2004). "A new model of information behaviour based on the search situation transition schema." *Information Research*, 10(1) paper 203. Retrieved from http://InformationR.net/ir/10-1/paper203.html.

Pharo, N., & Järvelin, K. (2006). "'Irrational' searchers and IR-rational researchers." *Journal of the American Society for Information Science and Technology*, 57(2), 222–232.

Pilerot, O. (2013). "A practice theoretical exploration of information sharing and trust in a dispersed community of design scholars." *Information Research*, 18(4) paper 595. Retrieved from http://InformationR.net/ir/18-4/paper595.html.

Pilerot, O., Hammarfelt, B., & Moring, C. (2017). "The many faces of practice theory in library and information studies." *Information Research*, 22(1), CoLIS paper 1602. Retrieved from http://InformationR.net/ir/22-1/colis/colis1602.html.

Rayward, W. B. (2008). "The march of the modern and the reconstitution of the world's knowledge apparatus: HG Wells, encyclopedism and the world brain." In W. B. Raward (Ed.), *European modernism and the information society: Informing the present, understanding the past* (pp. 223–239). Aldershot: Ashgate.

Rieder, B. (2009). "Democratizing search? From critique to society-oriented design." In K. Becker & F. Stadler (Eds), *Deep search: The politics of search and beyond Google* (pp. 133–151). Innsbruck: Studien Verlag.

Rieh, S. Y. (2004). "On the web at home: Information seeking and web searching in the home environment." *Journal of the American Society for Information Science and Technology*, 55(8), 743–753.

Rivano Eckerdal, J. (2012). "Information sources at play: The apparatus of knowledge production in contraceptive counselling." *Journal of Documentation*, 68(3), 278–298.

Rogers, R. (2013). *Digital methods*. Cambridge, MA.: The MIT Press.

Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., Gunter, B., Withey, R., Jamali, H.R., Moallem, T., Dobrowolski, T., & Tenopir, C. (2008). "The Google generation: The information behaviour of the researcher of the future." *Journal of Documentation*, 60(4), 290–310.

Saracevic, T. (2016). "The notion of relevance in information science: Everybody knows what relevance is. But, what is it?" *Synthesis lectures on information concepts, retrieval, and services*. San Rafael, Calif.: Morgan & Claypool Publishers.

Savolainen, R. (1995). "Everyday life information seeking: Approaching information seeking in the context of 'way of life'." *Library & Information Science Research*, 17(3), 259–294.

Savolainen, R. (2017). "Berrypicking and information foraging: Comparison of two theoretical frameworks for studying exploratory search." *Journal of Information Science*, 44(5), 580–593.

Schroeder, R. (2015). "Does Google shape what we know?" *Prometheus: Critical Studies in Innovation*, 32(2), 145–160.

Schroeder, R. (2018). *Social theory after the internet: Media, technology, and globalization*. London: UCL Press.

Schultheiß, S., Sünkler, S., & Lewandowski, D. (2018). "We still trust Google, but less than 10 years ago: An eye-tracking study." *Information Research*, 23(3) paper 799. Retrieved from http://InformationR.net/ir/23-3/paper799.html.

Singhal, A. (2001). "Modern information retrieval: A brief overview." *IEEE Data Eng. Bull*, 24(4), 35–43.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices." *Journal of Documentation*, 72(6), 990–1007.

Sundin, O., Haider, J., Andersson, C., Carlsson, H., & Kjellberg, S. (2017). "The search-ification of everyday life and the mundane-ification of search." *Journal of Documentation*, 73(2), 224–243.

Talja, S. & Nyce, J. M. (2015). "The problem with problematic situations: Differences between practices, tasks, and situations as units of analysis." *Library & Information Science Research*, 37(1), 61–67.

Taylor, R. S. (1968). "Question-negotiation and information seeking in libraries." *College and Research Libraries*, 29, 179–189.

Trias i Mansilla, A. & de la Rosa i Esteva, J.L. (2013). "Survey of social search from the perspectives of the village paradigm and online social networks." *Journal of Information Science*, 39(5), 688–707.

Tripodi, F. (2018). "Searching for alternative facts: Analyzing scriptural inference in conservative news practices." *Data & Society*. Retrieved from https://datasociety.net/output/searching-for-alternative-facts/.

Tuominen, K., Savolainen, R., & Talja, S. (2005). "Information literacy as a sociotechnical practice." *The Library Quarterly*, 75(3), 329–345.

Vaidhyanathan, S. (2011). *The Googlization of everything: (and why we should worry)*. Berkeley: University of California Press.

Vakkari, P. (2008). "Trends and approaches in information behaviour research." *Information Research*, 13(4) paper 361. Retrieved from http://InformationR.net/ir/13-4/paper361.html.

Van Couvering, E. (2008). "The history of the Internet search engine: Navigational media and the traffic commodity." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 177–206). Berlin, Heidelberg: Springer.

Van Dijck, J. (2010). "Search engines and the production of academic knowledge." *International Journal of Cultural Studies*, 13(6), 574–592.

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. New York: Oxford University Press.

Veinot, T. C. (2007). "'The eyes of the power company': Workplace information practices of a vault inspector." *The Library Quarterly*, 77(2), 157–179.

Waller, V. (2011). "Not just information: Who searches for what on the search engine Google?" *Journal of the American Society for Information Science and Technology*, 62(4), 761–775.

Waller, V. (2016). "Making knowledge machine-processable: Some implications of general semantic search." *Behaviour & Information Technology*, 35(19), 784–795.

White, R. W. (2016). *Interactions with search systems*. Cambridge: Cambridge University Press.

Williamson, K. (1998). "Discovered by chance: The role of incidental information acquisition in an ecological model of information use." *Library & Information Science Research*, 20(1), 23–40.

Wilson, T. D. (1994). "Information needs and uses: Fifty years of progress?" In B. C. Vickery (Ed.), *Fifty years of information progress: A Journal of Documentation review* (pp. 15–51). London: Aslib.

Wilson, T. D. (1999). "Models in information behaviour research." *Journal of Documentation*, 55(3), 249–270.

Wilson, T. D. (2010). "Fifty years of information behavior research." *Bulletin of the American Society for Information Science and Technology*, 36(3), 27–34.

Yeo, R. R. (2001). Enc*yclopaedic visions: Scientific dictionaries and Enlightenment culture*. Cambridge: Cambridge University Press.

Zavadski, A. & Toepfl, F. (2019). "Querying the Internet as a mnemonic practice: How search engines mediate four types of past events in Russia." *Media, Culture & Society*, 41(1), 21–37.

Zimmer, M. (2008). "The gaze of the perfect search engine: Google as an infrastructure of dataveillance." In A. Spink & M. Zimmer (Eds), *Web search: Multidisciplinary perspectives* (pp. 77–79). Berlin, Heidelberg: Springer.

Zimmer, M. (2010). "Web search studies: Multidisciplinary perspectives on web search engines." In J. Hunsinger, L. Klastrup & M. Allen (Eds.), *International Handbook of Internet Research* (pp. 507–521). London: Springer Science.

3

THE MATERIALITIES OF SEARCH

A Google search for "Pizza" in June 2018 from a Swedish IP number on a desktop computer resulted in over 1 billion (1,010,000,000) results retrieved in 0.61 seconds. We click our way through the results pages and the figure indicating retrieval time changes, but it is always under one second. On page 33, we still appear to have over 1 billion results to go through. Yet already on page 34 we reach the end and the figure changes to 334. We read:

> In order for you to get the most relevant results we have omitted some results that are similar to the 335 [sic!] already shown. If you want, you can redo the search and include the omitted results.

It seems unlikely that there would still be over 1 billion further results that are all similar to the 334 already shown. In Sweden alone there must be more than 330 pizza restaurants or delivery services with web pages. Yet, we redo the search as suggested by clicking on the link provided. Now Google tells us it has retrieved 1,060,000,000, that is an additional 50 million results, in just 0.6 seconds. Once again, we click through the results pages. This time we reach the end on page 70. Given that each results page shows 10 results and that the last page only had two, this means we have scrolled and clicked through a list displaying 692 results. This time, as we reach the last page, Google continues to state that it has retrieved over a billion results, yet we are offered access to just a tiny fragment of those. We change the display mode and ask Google to list 100 results on each page. Clicking through the pages and redoing the search after 440 results, this means we get a list of approximately 900 pages of relevance for the search term "pizza". The last one of those is the menu of a pizzeria in Illinois, USA. It is easy to understand why it is ranked very low, considering we are based in Northern Europe when we submit our query. The maximum number of results Google displays is in fact a thousand

DOI: 10.4324/9780429448546-3

(Lewandowski 2013, p. 148). Yet, how was this cut-off point selected? What would the next result have been? What is considered the least relevant result from other, less geographically sensitive, questions? On which data are the searches carried out and how is it processed to make these decisions possible? Bing, Yandex, and Baidu all work in the same way; they indicate the retrieval of hundreds of millions of results and then list a few hundred. Yet, only Google also adds information on its retrieval time. Clearly hundreds of millions or a billion results are not meaningful for humans, so why do web search engines bother telling us, just to then disappoint us with a few hundred links?

In Chapter Two, we presented selected approaches to understanding search and search engines that have developed over the last decades. Alongside research on information behaviour and information retrieval, we also emphasised work on the political, economic, and social structures shaping search engines, their workings or reasons for cultural dominance and some implications of these. What the present chapter wants to achieve is to return to some of the approaches presented in the previous chapter – on how we can understand searching for information online – in order to embed their insights into a discussion of how we conceptualise the materialities of search in contemporary society. In this way we bring together search and the tools used for searching, i.e. search engines, so as to open up a discussion of how these are increasingly invisible and powerful at the same time. Thus, this chapter elucidates conceptual devices relevant to understanding the dominance and constitutive embeddedness of search and search engines in society. To do this, we introduce three materially and technically consequential as well as theoretically and historically rich concepts – infrastructure, data, and algorithms. They are used to present some of the underlying technical structures of search engines and in doing so to continue the job of framing search as a sociomaterial practice that we embarked on in Chapter One.

Infrastructures

Talking about infrastructure means entering a strangely complicated area. There is the difficulty that it is a word used in everyday language and has a certain rather straightforward meaning for most people. Infrastructures might be technically complicated, but when we hear the word in relation to other infrastructures, for instance traffic or telecommunications, we have a good understanding of what is referred to, or at least so we think. It is the material, the hardware, the stuff that is arranged according to certain rules and that makes things work and makes activities possible: roads, rail tracks, cables, data centres and so on. A Google image search for *infrastructure* returns a collage of stylised images of near-carless motorway intersections and futuristic cityscapes with light streaks symbolising connections and movement. It looks nearly identical in Bing, Yandex, or any other search engine – clean, high-tech, and stereotypically futuristic. On the other hand, talk of *crumbling infrastructures* to be addressed with *large-scale infrastructure investments* is a staple in political discussion, and policymaking is full of infrastructure plans to address

disparate sets of issues, from a clean water supply to electronic communication, defence or space travel.

Various academic disciplines, professional groups, and industries have added qualifiers to mark their contribution to the understanding and creating of infrastructures: traffic infrastructure, financial infrastructure, institutional infrastructure, energy infrastructure, sustainable infrastructure, human resources infrastructure, computing infrastructure, to name just a few. In public and policy discourse, infrastructures are described in terms of *critical, vital* or *basic*; some are considered *soft* and others *hard*. What many infrastructures share is also an idea, however vague and diluted, that they are a common resource in some way or another, that they are regulated according to standards and– in some cases – that they are collectively owned and maintained. The "modern infrastructural ideal" (Plantin et al. 2018, p. 300) that dominated until the 1980s advanced this vision of government owned, large-scale investment in "publicly regulated monopolies" (ibid.). At least a form of collective ownership is the case for our most iconic infrastructures, roads, in many places also for sanitation and water supply and for a period it was the case for most rail or electricity networks. Actually, as Plantin and colleagues (ibid.) highlight, the early days of the internet, from the 1960s to the late 1980s were also characterised by this modern infrastructural ideal: "heavy government investment, sponsored first by the US Defense Department's Advanced Research Projects Agency, and then by the US National Science Foundation (NSF) in the public interest". Not only in the USA but also in Europe this was a dominant model: "The French Minitel system, introduced in 1980 and reaching 6.5 million French citizens by 1990 (far more than any commercial networking service), represents perhaps the most complete realization of the computer utility model as a public good" (ibid.). Since then, infrastructures of all kinds have been broken up and privatised, sometimes left to market forces and deregulated. Still, this has done nothing to diminish their significance, materially, symbolically, or discursively.

Critical infrastructure

The notion *critical infrastructure* implies that certain infrastructures support societal needs that are so basic that they need to be designated and protected in particular ways. That is, even if they are not collectively owned or nationalised, they are still tightly regulated and, in an emergency, they can be controlled by a surrogate of the state so as to ensure the working of vital functions in society, and control over them is thus crucial. The World Bank has identified the "infrastructure gap", threatening financing of infrastructure development throughout the world, and a search for the term *infrastructure* in its various databases yields several hundred thousand results. China provides so-called infrastructure loans worth billions of dollars to countries in sub-Saharan Africa for infrastructure projects considered critical for economic development, ranging from dams, to rail tracks or telecommunications projects (Foster et al. 2009), thus expanding and solidifying China's sphere of geo-political influence, and similar investments are underway in

South American Countries (UNASUR n.d.). In the European Union, a dense network exists of documents, directives, working and implementation groups and the like, all dedicated to the identification and securing of critical infrastructures on national and cross-national levels (European Commission 2018). In the United States, the Department of Homeland Security identifies 16 sectors as providing critical infrastructures essential to maintaining the functioning of society – everything from healthcare and financial services to, of course, information technology.

Paul Edwards highlights pointedly, "Given the heterogeneous character of systems and institutions referenced by the term, perhaps 'infrastructure' is best defined negatively, as those systems without which contemporary societies cannot function" (Edwards 2003, p. 187). In this sense, internet technology and communication have developed into what can be considered critical infrastructures for large sections of contemporary society (see also Sandvig 2013). However, in most regards, it is hard to pinpoint where exactly the infrastructure ends and a service running on top of that infrastructure starts. Which part of the internet is infrastructure, which part is critical, and which part is simply "using" this infrastructure in order to provide a service? Moreover, which of these services are in themselves so pervasive and at the same time critical for other activities that they need to be considered to be infrastructures in their own right? These demarcations are fluid and ever changing and they have to be.

The privatisation and fragmentation of many infrastructural arrangements, especially in the information technology sector, further complicate these delimitations. In this context, it is thus interesting that the European Union, in the Directive on Security of Network and Information Systems (NIS Directive) from 2016, describes certain information technology providers precisely in terms of critical infrastructures, specifically providers of three types of digital services are singled out: cloud computing, online marketplaces, and online search engines. They are also put alongside drinking water supply, financial and banking services, transport, or energy as essential for the functioning of society. The identification of search engines as critical infrastructure by the European Commission (Council Directive 2016/1148/EU, p. 15) also includes a definition:

> "online search engine" means a digital service that allows users to perform searches of, in principle, all websites or websites in a particular language on the basis of a query on any subject in the form of a keyword, phrase or other input, and returns links in which information related to the requested content can be found.

Infrastructures are here political objects, fluid and strategic, and at the same time profoundly consequential. Positioning online search engines as an infrastructure, and in particular as a critical infrastructure, underscores their strategic importance and their indispensability for accessing information; in other words, doing so explicitly acknowledges their wide-reaching control over information access and thus over information. It is also a further affirmation of their embedded invisibility

in society and everyday life. Their definition as merely providing links to information shows, however, the difficulties involved in turning search engines into a targetable object for policymaking. In the light of, for instance, Google Search's move towards providing answers rather than links to documents containing the answers, and more generally its transformation into what could be called a "suggest engine", a development we mention in Chapter One, this is a somewhat limited and limiting definition. Moreover, it is a superficial one in that it really only concerns the surface, but leaves out what some consider to be the most valuable part of contemporary search engines, their index. We will return to the question of the index, an acutely material, very fundamental yet obscure part of online search engines, in the subsequent section. However, first we need to conceptualise infrastructure theoretically.

Infrastructure as a conceptual device

While the internet and with it myriads of variously interdependent and interconnected digital services developed in a way that made their depiction and regulation as a critical infrastructure not only possible, but also lets such a depiction appear self-evident, the term infrastructure also developed into a theoretically rich concept in the social sciences and the humanities. Some of this development is introspective. Here, it concerns, on the one hand, the material conditions, not least libraries, archives, and publishing structures, that the social sciences and humanities have traditionally been invested in and, on the other hand, new methods and areas of research opening up with digitisation (e.g. Burdick et al. 2012).

A substantial part of the theoretical development around the notion of infrastructure stems from investigations into the distinctive qualities of material and organisational structures shaping society, and specifically those implicated in information and knowledge production. This is to a large degree facilitated through a sociomaterial perspective, as developed in science and technology studies (Bowker et al. 2009), organisational studies (Orlikowski 2007; Orlikowski & Scott 2008), but also social informatics (Kling 2000), and goes back to early studies of computerisation in the 1980s and 1990s.

> A Web search conducted with the Google search engine is sociomaterial "all the way down," entailing computer code written and updated by software engineers, executing on computers (configured with particular hardware and software elements which were designed and built by computer engineers and production workers)

explain Wanda Orlikowski and Susan Scott (2008, p. 465) not least since, as they emphasise (ibid.), its

> operation depends on the millions of people who use computers to create and update Web pages everyday, and the millions of people around the world who enter particular search criteria into their Web browsers running on still other computers designed and built by yet other people, and so on.

The concept of infrastructure was shaped and re-imagined in ways that have turned it into a powerful yet subtle tool for analysing and understanding the various mechanics for dealing with information in very different settings. In the early 2000s, these approaches coalesced into the field of "infrastructure studies" (Edwards et al. 2009). To reiterate, infrastructure is not reserved for discussing information, nor can information infrastructure simply be likened to information and communication technologies (ICTs). Yet, a substantial part of the conceptual apparatus attached to discussions of infrastructure has developed in close relation to thinking about digital technology and what it does with how we deal with information. Likely this happened because some of the old infrastructures only became visible when newer ones challenged them. The notion of information infrastructure emerged describing arrangements of material relations implicated in how information is produced, circulated, used or otherwise encountered in specific settings (Bowker 1994, 1996). However, by shifting the focus onto the function fulfilled by an infrastructure, rather than onto a specific technical solution in a certain moment in time, these settings necessarily also include non- or pre-digital phenomena (Edwards 2003). In the words of Paul Edwards (2003, p. 207): "Seen as infrastructure, information systems are ways to handle the functional problems of information storage, transfer, access, and retrieval", and this clearly also includes institutions such as libraries, archives, museums, the book publishing sector, and others (see also Mattern 2014). Undoubtedly, the most influential account for the information sciences, in this context, can be found in the work of Susan Leigh Star together with Karen Ruhleder (2015 [1996]) and also Geoffrey Bowker (Bowker & Star 1999; Star & Bowker 2010). Rather than thinking about infrastructures as immutable, technical, and independent of use, they conceptualise infrastructures as situated, relational, temporally structured, as well as – and this is particularly relevant for our account of search and search engines in everyday life – deeply anchored in practice.

Search as information infrastructure

An information infrastructure might appear as a technical solution bound up with certain activities in a direct way – talking on the phone, cooperating online in a work setting, querying a database, looking up facts via a search engine. Yet, the sometimes all-encompassing grip certain infrastructures have on everyday life and work practices and their significance as political, economic, and cultural objects in relation to this grip, requires us to look more closely into how, where, when, and to what effect this interlacing of activities and material structures takes place. Susan Leigh Star and Karen Ruhleder (2015[1996]) remind us that infrastructures are not stable; rather, they write, "infrastructure is something that emerges for people in practice, connected to activities and structures" (ibid., p. 379). Infrastructures are here transparent and embedded, that is they are "'sunk' into, inside of, other structures, social arrangements and technologies" (ibid., p. 380). Their use is naturalised; at first encounter, infrastructures are new, strange and thus visible. They are being made transparent as we learn to use them more and more proficiently until we take them for granted and they merge into the background. Technically,

infrastructures incorporate standards. They are built on top of other infrastructures and they change and evolve incrementally (ibid., p. 381). Put simply, the internet emerged on top of telephone lines, and search engines emerged on top of the internet. The most frequently mentioned feature of infrastructures concerns how we notice them despite their characteristic embeddedness and invisibility; we notice them on breakdown. "The normally invisible quality of working infrastructure becomes visible when it breaks; the server is down, the bridge washes out, there is a power blackout", write Star and Ruhleder (2015[1996], p. 381) akin to Edward's (2003) notion of infrastructures being best defined negatively.

If we return to the way in which regulators attempt to define online search engines as critical infrastructures, then this also has to be understood in the context of how closely they are integrated with everyday life. The use of online search engines is in itself part of the infrastructure of everyday life. However, this does not happen in general, abstract terms, but in relation to specific social practices. It is thus also in these very relations where an investigation of search as infrastructure has to start from.

Rethinking visibility and breakdown

Two of the characteristics of infrastructures introduced above require some more reflection in relation to search engines, namely the notion that they become visible on breakdown and that they are "learned as part of membership" (Star and Ruhleder 2015[1996], p. 380), which in turn also has implications for when and how an infrastructure can be noticed. Leigh Star and Karen Ruhleder (2015 [1996]) draw on the work of Lave and Wenger (1991) to stress the importance of so-called *communities of practice*, where an infrastructure is negotiated and shaped in relation to the community's goals and where it is thus normalised and made transparent. Applying such an understanding to society's dominant online search engines as they operate today has its difficulties. On the one hand, it describes the situation very well and it is easy to accept that this is how transparency is created and inversely how an infrastructure is especially noticeable when we are not (yet) used to it. On the other hand, however, if we replace the unspecific term online search engine with the name of the most dominant actually existing ones, Google and, in some distance, Bing, Yahoo, Baidu, or Yandex, then the issues of invisibility on breakdown and learning as part of a community of practice become considerably more complex. When do they actually ever break down; which community are they a part of, and how can we even meaningfully pinpoint them?

Contemporary online search engines are not only invisible in the sense that they go unnoticed when they work as they should, but are invisible also in a number of other ways. Firstly, their digital materiality is multiple. Secondly, they are dispersed and mobile. Thirdly, as they are used across communities and across practices, search engines are rarely strange in any meaningful way. Finally, rather than understanding breakdown solely as a technical occurrence, we need to position it – also – at the interlacing of practices and material structures, in line with the

conceptualisation of infrastructures as situated, embedded and emergent through use. Let us proceed to unpack these propositions.

Firstly, a search engine is not stable, but constantly "in the making". Search engines change through their use. Their index constantly changes as the web grows and content gets updated, deleted, and added. Likewise, their algorithms for indexing, selecting, ordering, and displaying are constantly updated through use. They grow and learn. Search engines draw on use and user data in order to improve their relevance for the individual searcher, but also for the collective development of the search engine. The search engine of one user is always different from that of another, however slightly; they are elusive entities that can only be perceived meaningfully when they are used, in their effect, and in this sense, they can only ever be fixed temporarily.

Secondly, online search engines are a distributed, dispersed, and acutely mobile information infrastructure. A search engine is not stuck within one device, reserved for a specific interface or tied to one practice. Search engines spread out and meander; they are everywhere, in phones, tablets, computers, cars, watches, and so on. Their interfaces are text-, image- or voice-based and adapt according to a device's specifications. The so-called internet of things embeds them into ever more objects, places, and practices, and buries them ever deeper inside those. Often, we use search engines without any indication whatsoever that this is what we are doing. This is not unique for online search engines. In today's society, one is rarely outside the realm of digital information infrastructures. Even if we ourselves are not online or using digital tools, everything and everyone around us usually is and does, from the librarian or the bus driver to the cashier, from the photocopier, or the CCTV camera, to the car passing by.

Thirdly, search engines form practices and are naturalised far beyond the reach of specific communities of practice. Certainly, Google has different meanings in different communities and in relation to different practices, and so have Baidu, Yandex, or any other culturally dominant search engine. Using Google for a school exercise or to find out about a tourist attraction, searching on Baidu for an opening time, or using Yandex to navigate to a news site are different from using them, for instance, in the context of a work problem, related to a health issue or to establish and judge the position of a political organisation. All these searches are carried out in certain contexts, in certain roles, as a student in relation to the school's requirements and rules, as a patient in relation to the care structures, as citizens of a particular municipality, as a tourist, and so on. Examples abound, but the point is that, while all these uses relate to different sets of practices and make the use of the search engine part of this practice according to rules and norms that are specific to the respective community or purpose, they all use the same search engine, on very similar devices. The search engine is thus never strange, never new, and in this sense, it is never truly visible. Almost nobody is ever a beginner, at least not in a conscious way where learning to use the tools is part of how one becomes part of a community. The doctor, the nurse, the patient – all use Google, as does the teacher, the pupil, and the librarian. It is safe to assume that in many

countries most people born after the year 2000 have never *not* known Google, even if it is largely invisible to them. The school setting constitutes an interesting and to some extent exceptional case, as it so clearly provides Google Search with a different context, one in which critical evaluation of information is crucial and also more consciously reflected on than in other situations. Despite the difficulties that schools have in translating Google Search to the practices of school-based learning (Sundin 2015), the school often tries to make visible what in everyday life is invisible (Andersson 2017; Sundin & Carlsson 2016). We will discuss this further in Chapter Five.

Fourth and finally, the question is when do online search engines not work? Understood in a purely technical sense, it happens of course, and it takes little imagination to understand how the disruption arising from such a technical breakdown would make their critical infrastructural functioning acutely noticeable. This is also why a body like the European Union designates search engines as critical infrastructures. However, it hardly ever occurs and if it does, it happens for very short periods of time, which people might not even notice. Having said that, the notion of infrastructure as situated, relational and emergent in use also means that breakdown occurs not just by itself in the technical, material structure, but also in relation to practices and situations (Haider 2016; Pilerot 2014). For instance, in some situations carrying out a search is simply not considered to be socially acceptable. Other situations require a restricted use of search technology in order to keep certain information hidden. Ethical considerations might make using large commercial search engines problematic, or the political conditions in a place can result in the use of a particular search engine being impossible. A lack of skills could make a certain type of search impossible or difficult, as could the interface's affordances. All these situations would amount to a type of infrastructural breakdown beyond pure technical malfunctioning, a breakdown situated at the point where practices and various material conditions constitute each other. We will return to these breakdowns in greater detail in Chapter Four. To sum up, infrastructure is made meaningful in practices and communities around these practices. However, as regards search engines, in most cases the same technology is tied up in a large variety of practices, across communities and in different stages of life. They all shape this information infrastructure together and thus – we must assume, and this is what is at stake in Chapter Four – they also become implicated in each other.

Platformisation of the information infrastructure

Thinking about search and search engines in terms of infrastructure helps us to trace their foundational envelopment in contemporary everyday life. However, it also obfuscates that when we talk about search engines and how we use them in everyday life, we rarely talk about them in an abstract sense. Rather, we refer to – depending on country – Google Search, Baidu, or Yandex and very rarely others and if so then most often in relation to Google. Furthermore, Google Search as well as Baidu or Yandex are much more than "just" search engines. They exist within arrangements of numerous integrated services, applications, and programs

working together and which can latch onto other services that adhere to certain technical protocols. These are constantly growing by adding new services to their platforms, by expanding the reach of existing ones or by acquiring start-ups. They are advertisement companies, they are media companies, they are developer companies, they traffic in user data, and so on. In line with their profit interest and business model, their aim is for users to be become locked into their universe of services rather than into that of a competitor. The question is whether the notion of infrastructure alone is in fact sufficient for us to fathom the role and function of these dominant search engines, where economic power, social control, and material structures work together to concentrate control over information and all this implies in the hands of ever fewer multinational corporate actors.

To describe the specific position of dominant information intermediaries on the contemporary internet the concept of platform has emerged. As Tarleton Gillespie (2010) develops it, the notion draws together a number of discursive arenas, each with their own understanding yet working together to advance a broader meaning, one that speaks to different aspects and different constituents, developers and programmers, the business community, but also users and content producers. A purely computational understanding, positioning platforms as "an infrastructure that supports the design and use of particular applications, be they computer hardware, operating systems, gaming devices, mobile devices or digital disc formats" (Gillespie 2010, p. 349) has been left behind in favour of broader, also more metaphoric, meanings. These work to unite otherwise conflicting functions or to address different constituents, in software development, business, marketing, cultural production and so on. Importantly for us, the notion of platform has begun to capture the specificities of the commercial formations governing the contemporary internet, an internet largely dominated by GAFAM (Google, Apple, Facebook, Amazon, and Microsoft) in the West and BAT (Baidu, Alibaba, Tencent), the so-called "three kingdoms of the Chinese Internet" (Keane 2017). The concept of *multi-sided platform* describes companies whose business model is based on enabling interactions between at least two partners, for instance users and advertisers, software providers and consumers and so on (cf. Hagiu & Wright 2015). Bernhard Rieder and Guillaume Sire (2014) draw on this conceptualisation to describe Google Search as a three-sided market, and in broad strokes this clearly also applies to Baidu, Bing or Yandex.

> On one side, Internet users query the engine to find information, entertainment, and so on. On a second side, Google indexes "content providers" that want users to reach their websites. On the third side, advertisers are trying to attract visitors *beyond* the traffic received from "organic" results. Google subsidizes two of the three sides and charges the third: Internet users search the Web for free; content providers are charged neither for getting indexed nor for the traffic they receive from organic results; advertisers, however, pay for every click and thereby finance the platform.
>
> *Rieder & Sire 2014, p. 200*

Many times, of course, content providers and advertisers are the same people. The term *organic results* refers to the search results that are not visibly marked as advertisements; they are not paid for or ranked due to payment. In the study of contemporary information intermediaries, the characteristics of platform and infrastructure cannot meaningfully be teased apart. Rather, as Plantin and colleagues suggest, and we agree: "Digital technologies have made possible a 'platformization' of infrastructure and an 'infrastructuralization' of platforms" (Plantin et al. 2018, p. 295). Thus, understanding search engines, as infrastructures, but doing it through the notion of platform, or vice-versa, opens up for addressing search engines as deeply embedded in everyday life and shaped by, what Astrid Mager (2012) calls, an algorithmic ideology infused by the spirit of capitalism.

Data: Index, queries, and user data

Developing information behaviour or retrieval research in a situation pervaded by a certain type of general-purpose search engine, requires a basic understanding of how some central elements of contemporary search infrastructure are constituted. Accordingly, we need to go into some detail about how search engines and more specifically Google Search actually technically function. Queries, index, user data, algorithms as well as, of course, users and their practices are in a sociomaterial perspective all regarded as actors that constitute each other. We start with some reflection on the data involved and then continue by discussing algorithms and their workings, always aiming at striking a balance between conceptual considerations and technical specifics. Our notion of data is informed by the much-cited saying that data is never *raw*, but always already *cooked* (Bowker 2005; Gitelman & Jackson 2013). In other words, before data can even be compiled as data, it needs to be imagined and specified. These specifications exist within and in relation to various societal and material conditions (see also Kitchin 2014) that are then intrinsic in how things are turned into data and how data is collected, organised and used.

It might be tempting to say that search engines are software programs implementing algorithms to enable term matching (and to display the results on an interface). However, this neglects the most central part of a search engine, namely the material that is being sifted through when a search is carried out. This collection of web pages, the so-called index, is easily missed. It is hardly ever discussed and never seen. Yet, without it search engines cannot work. Control of a well-developed, updated index is crucial for search engines to function, or indeed to have any function at all. A web index is extremely costly to acquire and maintain, not to mention the computing resources, bandwidth, storage, energy consumption and actual physical and human infrastructure that its upkeep requires. This has serious implications for the concentration of power that characterises the contemporary information landscape. Dirk Lewandowski, one of the few who has repeatedly called attention to the societal and also political significance of the search engine index, asserts: "New search engines fail at the creation of the index" (Lewandowski 2013, p. 158, our translation; see also Lewandowski 2014).

Web search engines are sets of interacting software programs and continuously updated collections of data sets that enable the matching of search queries to a corpus consisting of copies of web pages and other web material and the reporting back of presence (or absence) of the query to a searcher, either a person or another computer program. They handle and produce data in digital form, and some of this data is in fact part of the search engine. Put bluntly, search engines need stuff to search *on*, information, data, material, in the form of web pages, images, text, and so on. They also need stuff to search *with*, metadata of different sorts that explain and situate the material to be searched and how or whether it should be searched. However, in addition to data to search on and data to search with, the use of contemporary search engines also generates data, massive amounts of data in fact. The search engine then uses the data produced in this way to improve its algorithms, its index, and also its understanding of a specific user, or group of users, and thus of what types of results the person – and others like them – might find relevant.

On a conceptual level, it is not an easy task to distinguish between the different categories of data that search engines work with and which they produce, from each other or from other parts of the search engine, for that matter. They are created in relation to each other and are thus also inscribed into each other. Nevertheless, in order to enable a more precise understanding of the workings of search engines and their significance in contemporary everyday life, a focus on some of the distinct qualities of specific types of data and the processing they undergo is useful. In the following, this is done by focusing on two types: Firstly, the data that is being sifted through, the index, and secondly, various kinds of user generated data that are produced and collected to typify users and queries and which feed back into the engine. In addition, as a bridge between these two – index and user data – some reflections on the data underlying autocomplete suggestions will be provided. All these are part of the search engine. They are this not in any metaphorical way, but they are – admittedly to different degrees and in different ways – intrinsic to the functioning of search engines. They are materially and to a degree also legally under the control of a search engine and its owners. Admittedly, this is a selective focus on just a few and very specific types of data that are at work in search engines. Yet, such a focus works to the effect that it highlights particularly well the situatedness of data in the case of contemporary general-purpose web search engines. This, in turn, contributes to establishing a base from which to consider some wider implications in relation to the purpose of this book, which is to provide a reflection on the meaning and implications of the ubiquity of search engines in everyday life.

Index

It is a platitude to say that we do not know the size of the web. How many web pages, how many files, how many gigabytes, how much data are there at any point in time? This is not surprising, of course. There is no central organisation doing the measuring and it changes continuously. However, since a search on the live web is

not realistically feasible and, were it possible, would be far too slow, web search engines create copies of as much of the web as they can to do the searches on. In many respects and for most everyday uses, the web is actually equal to the partial, data-processed copies that are stored on Google's, Microsoft's (Bing), Yandex', or Baidu's hard drives, as they are warehoused in gigantic, energy-devouring server halls, preferably located in cool climates or close to a water source. Yet, we also know very little about these databases.

It is commonly accepted that Google's database, its index, is the most comprehensive, most frequently updated and generally the best-developed one. Regarding its size, Google very vaguely declares, "[t]he Google Search index contains hundreds of billions of webpages and is well over 100,000,000 gigabytes in size". (Google n.d. c) This statement does not really say much about the index, other than that it is huge beyond comprehension and that Google has control over it and over how it is processed. Yet, it leads to two relevant questions. How do web search engines go about finding and selecting the material to be included in this gigantic database? And how does the creation and matching of entries for the index function in practice?

Lewandowski (2013, pp. 144–150) distinguishes between three methods for content acquisition by search engines, namely crawling, the import of feeds and the retrieval of structured data from databases. Crawling, which is the main strategy, takes advantage of the web graph. A software tool called a crawler, sometimes also spider or robot, automatically finds and downloads material from the web by following links. "*[C]rawling* refers to the (un-) focused retrieval of a collection of Web pages with the purpose of making them available for subsequent analysis, cataloguing and indexing of a search engine" write Paolo Ferragina and Rossano Venturini (2013, p. 120) thus stressing the close relation between acquisition of content and its processing. Crawlers constantly retrieve new material, but also update material that is already downloaded or verify whether material still exists and delete it from the index if needed (Lewandowski 2013, p. 145). A second method for content acquisition consists of the downloading of feeds. Here the search engine *pings,* that is it sends a request to the providers of web feeds for documents, and if new ones exist, they are downloaded in structured form, a method that is, for instance, used to download and update entire product catalogues (ibid., p. 147). Finally, the third main method search engines use to populate their index is the direct integration of structured data from databases (ibid., p. 147). Additionally, search engine providers also themselves produce data or material to be searched. This is the case with other services owned by them such as, for instance, Maps in the case of Google, and they also host user produced material on their own servers, such as for instance in the case of YouTube (ibid., p. 36). Obviously, it is impossible to produce an exact copy of the web, and what we get is a patchy approximation. Some content is not included because its providers actively seek not to have it indexed. This is technically straightforward and easy to achieve. Other content, as Lewandowski (ibid.) also highlights, might be included in the index, but does not show up in response to a search because it might be against the law or

a policy. The latter would for instance be the case with content, personal data, which is removed from the search results appearing in response to certain keywords, under the European Union rule known as "The Right to be Forgotten" (Lindsköld 2018). The information is still available and also indexed, but it is delinked and cannot (easily) be accessed through searching for it.

Three interesting terms should be mentioned in relation to how a search engine index is populated and the significance of this largely invisible process for the cultural meaning of the internet, namely *surface web, deep web*, and *dark web*. The first, surface web, is used to refer to the portion of the web that can be found by means of general-purpose web search engines. The deep web, also called invisible web, is the part that cannot. It is content which is "hidden" behind forms, not linked to from anywhere, technically difficult to process, or not findable for good reasons, such as paywalls, safety, integrity and so on as would be the case with webmail or online banking (Madhavan et al. 2008; Lewandowksi 2018, pp. 251–262). The web search engine constitutes the boundary between the two and necessarily this is a boundary that moves. It could also be argued that in this categorisation, the Internet Archive, a non-profit collection of diverse public domain materials and most famously a huge archive of old websites, is part of the deep web. In an ironic twist, the internet archive cannot be googled, or searched with any other general-purpose web search engine, but only by means of its own search engine, the Wayback Machine. Quite suitably the memory of the web is fragmentary and pressed down below the surface. Nevertheless, the content that can be found with general purpose web search engines today is by far more comprehensive and also more varied than it was when the terms were first coined in early 2000 (Bergman 2001). The distinction between surface and deep web describes a technical boundary and originally it did not imply anything sinister. The last term in the list, dark web, however is different. It is generally taken to refer to illegal content or otherwise dubious, secretive activities hidden from view, mostly in the deep web but not necessarily so. Definitions abound and in particular its relation to the so-called *darknet*, the role of cryptography or peer-to-peer technology is part of various demarcations. This is beyond our scope. However, notably, the notion of the dark web has come to be mixed up with the deep web (Lewandowski 2018, p. 253), probably in allusion to conspiracy theories of the deep state or similar narratives of hidden power structures. The technical processes of how web search engines create their index and how they then make certain content accessible or not becomes thus part of the cultural narrative of secretive information structures beyond mainstream society's control.

Google itself explains the role of the search engine index by means of a puzzling analogy: "It's like the index in the back of a book — with an entry for every word seen on every web page we index. When we index a web page, we add it to the entries for all of the words it contains" (Google n.d. d). It is somewhat ironic that Google, who with Google Books has made the traditional book index close to irrelevant, should use precisely the book index as an analogy to explain its own index. A book index is a bibliographic device that contains *selected* terms in

alphabetic order, which the authors (or editors) deem relevant as a way to access the material in the book. The analogy draws attention to the fact that even a simple full text search in Google and of course other search engines does not actually sift through the copies of the webpages as they are, but these have been indexed, in other words prepared according to certain rules. The indexing phase concludes by "the construction of a data structure called an *inverted list*, which is the backbone of algorithms answering the user queries" (Ferragina & Venturini 2013, p. 125). The terms in such an inverted list are assigned weights, depending not least on how frequent and common or rare and specific they are, not just in the document, but in the collection (Spärck Jones 1972). In a book index, judgements will have been made as to what to include or exclude, which terms to refer to other related terms, which terms should have subcategories or narrower terms, and so on. Likewise, in a search engine's index, although all terms can be included, choices will still have been made concerning, for instance, how terms relate to each other, to terms in other documents, or more generally which information about a term or a document is worth preserving and in what form and so on. These decisions – even if they are automated and occur at high levels of abstraction – are informed by specific conditions, including, as has been highlighted repeatedly, the business model and socio-political circumstances of search engines. But, less obviously, they are also informed by established principles of knowledge organisation. These principles are, as we know (Bowker 1994; Bowker & Star 1999; Olson 2001), necessarily shaped in relation to specific cultural settings and their possibilities and constraints. As such, they also emerge from and are constitutive of various power relations and accordingly subject to change over time.

While considerable attention is being paid to search engine algorithms, the index and the principles it incorporates are hardly ever addressed. Plainly, this is because researchers cannot access the index. And, while algorithms are equally inaccessible, at least, since by definition they are output-oriented, there is a sense that their effects can be explored. Ferragina and Venturini (2018, p. 125), in the quote above, call the index's data structure the "backbone of algorithms", thus emphasising how in effect the index and the algorithms doing the retrieving and ordering of results are different sides of the same coin. One of the few indications we get that can help us work out just a minor aspect of how search engines actually relate terms to each other is through the *autosuggest* function. Here, the relationship between terms is established on how they co-occur in different searches. From these co-occurrences relationships between different documents (websites) are extrapolated. Often these appear reasonable in a common-sense kind of way, aiding searchers to formulate their queries and in this respect, they work very well. Yet, the connections established are consequential for the representation of issues. They steer users in certain directions, thus potentially solidifying or challenging knowledge claims even before the search results are even retrieved, let alone ordered (Haider 2016). And of course, for a business or an organisation showing up or not as a suggested term in autosuggest has financial ramifications. Sometimes, autocomplete suggestions so blatantly break with established norms and conceptions of what should be considered the established knowledge of an issue that this

makes palpable the authority held by these term relationships, but mostly they work without us paying much attention to their intrinsic power. This happens at an individual level and occasionally also at a collective level, such as when the media report sexist, racist, or in other ways discriminatory or problematic combinations of suggested search terms. To draw on the above discussed notion that infrastructures become visible on breakdown, it can be suggested that these moments when autocomplete suggestions are noticed because they infringe our value system or accepted perception of an issue are moments of infrastructural breakdown. As we touched upon in Chapter Two, various discriminatory practices and structures shaping society are also formative of search engines and vice-versa, as, for instance, Safiya Noble (2018) elucidates in great detail. Yet, on a personal level, it is exactly in these moments of infrastructural breakdown that these structures and how they operate in search engines become perceptible. The image of breakdown is not one that elicits a malfunctioning search engine, but one that prompts an interruption of the perception of the search engines as valueless and neutral. The infrastructure stops working frictionless (for us), and we notice not only the infrastructure in question, but how deeply mixed up we are in its functioning. "Google acts in the gap between the user's query and the content", write Deirdre Mulligan and Daniel Griffin (2018, p. 569). An occasion of this gap literally materialises, it could be argued, in the shape of the unassuming box into which a query is entered. The search box provides a crack through which we can catch a glimpse of the search engine's inner workings. Further on in this chapter, we develop this notion of breakdown in relation to relevance.

User data

Users create data each time they carry out a search. They enter a query, most often by typing search terms, but increasingly by posing a question aloud or by submitting an image. Users further produce data by then clicking on a result link, or by modifying their search and submitting it anew, by clicking on a recommended related search, by choosing a suggested search term that pops up in the query box, and so on. In other words, data is produced through the interaction of people with the search engine. This data is then stored in logs of user searches and drawn on as a form of feedback in order to improve the search engine. Michael Zimmer (2008, p. 91) likens Google specifically, but this can also be translated to other dominant web search engines, to "a black hole, /…/ using its gravitational forces to pull as much information about its users into its domain as possible". He maintains: "By monitoring and aggregating the results of every Web search performed, every image result clicked, every Website bookmarked, or every page visited with the Toolbar, Google has created [a] sophisticated infrastructure of dataveillance". Since today's dominant search engines operate as multi-sided platforms, it is not only interaction with the search engine itself that creates user data. Google, or more precisely Alphabet Inc., its parent company, offers a huge number of services, such as Gmail, YouTube, Calendar, Maps, and many more. Everything you do on these services is producing data for Google. Google also gets data from websites that use

services provided by Google, such as AdSense, DoubleClick, and Google Analytics. Ibrahim Altaweel, Nathaniel Good, and Chris Jay Hoofnagle (2015) show how users who visit the 100 most popular sites will have thousands of so-called cookies, i.e. small programs used to track certain interactions, saved on their computers. Furthermore, they show how it is specifically Google that dominates this cookie market and Google cookies exist on the vast majority of the web's most popular sites. As a consequence, for instance, when you search for something in Google, adverts keep popping up when you are on other websites or social media. This is part of a dense network of economic interdependencies where the search engine as a multi-sided platform reaches into all kinds of other online activities, postulating its ubiquity in everyday life and intrinsic control over our online whereabouts in very literal ways.

Felix Stadler and Christine Mayer (2009) provide a useful distinction for understanding different types of user data: a *knowledge profile*, a *social profile*, and an *embodied person*. The knowledge profile consists of what you as a user search for, what links you have followed, and where you have been. Other data are language used and which Google top domain is used. The social profile constitutes data that captures interaction with others in social networks, mail, and so forth. Not least important is the embodied person that concerns data where you live and other data in relation to location. Google themselves distinguishes three types of user data they collect (Google n.d. d): "things that you do" (e.g. search, watch, click etc.), "things that you create" (e.g. contacts, photos, calendar notes) and "things that make you 'you'" (e.g. name, gender, country, address). Google also offers help for users to get an overview of user data they generate and to some extent decide what data Google can link to each user account.

Stadler and Mayer (2009, p. 99), suggest that user data could be referred to as a second index compared to the one that consists of data about the web: "This one is not about the world's information, but about the world's users of information". As Stadler and Mayer maintain, the enormous amounts of data have the obvious advantage of helping Google to contextualise its search results on an individual level. However, they also stress the troublesome challenges all this data poses for the sake of privacy and integrity. As we showed in Chapter Two, searches are definitely personalised in terms of geo-location of the searcher (Kliman-Silver et al. 2015) and while this makes a lot of sense if we consider the billions of results a simple search for pizza generates, this can potentially also be highly problematic. A place is not just a place, but there are always assumptions about what type of people are residents of a certain town or area, and data to describe a location often includes average income, political leanings and so on. There are of course also other personalisations involved, depending on the topic and type of search, such as someone's search history. Yet these personalisations seem to have little impact on the organic search results (compare Hannak et al. 2013; Dutton et al. 2017). It also appears uncertain as to whether Google or other search engines use the type of demographic data that is often tied to geo-location for personalising the order of organic search results (see also Lewandowski 2018, p. 124). Having said that,

adverts are clearly heavily personalised and here various types of demographic data is processed to infer information about the individual user. In addition, as William Dutton and colleagues (2017, p. 149) show "autocomplete suggestions become more specific and personalized based on the semantic search histories of each user".

Algorithms

Considerable attention has been paid to algorithms in recent years, to the point that algorithm has almost become somewhat of a catchphrase in the humanities and social sciences, culminating in the field of critical algorithm studies (see also Bucher 2018). Yet algorithms are nothing new, nor have they come into being with social media and search engines. In fact, we can trace the use of algorithms back to Mesopotamia and Ancient Egypt (Ausiello 2013). Algorithms contain the rules for how an input is to be turned into an output. A textbook definition goes as follows: "an algorithm is any well-defined computational procedure that takes some value, or set of values, as input and produces some values, or set of values, as output" (Cormen et al. 2009, p. 5). This type of computer-science definition is relevant for us, yet we also need to consider that, as Taina Bucher (2018, p. 19) emphasises, "Algorithms exist and operate on multiple levels". They "can be magical *and* concrete, good *and* bad, technical *and* social", she reminds us. "Algorithms do not merely *have* power and politics; they are fundamentally productive of new ways of ordering the world as part of a much wider network of relations and practices" (ibid., p. 20).

Ordering and ranking

Algorithms have necessarily always been fundamental elements underpinning the working of information retrieval systems, and they are implicated across the entire process, from the crawler retrieving the webpages to the index being structured and to the search results being ordered and displayed. In the present section, our focus is on the latter. In all computerised systems for search, the search results are ranked and presented according to some kind of principle, such as chronology, subject, time of acquisition, or author. The huge number of web pages, the immense amount of potentially relevant information, has made it necessary to develop new strategies to rank search results in different ways than before. The major changes that search engines and social media have brought to how algorithms judge relevance is ranking based on popularity and, later, a personalisation of social media feeds and search engine results. The working of these algorithms is one of the bases for the competition between different commercial actors. For example, one of the most important reasons for Google to advance to the position of dominant search engine is said to be the implementation of the PageRank algorithm.

The algorithms of search engines change constantly. Finding out about the algorithms of commercial web search engines is only possible to some extent by leaning on the companies themselves. Some of the workings are made public,

not least under patent law, such as the PageRank algorithm (Brin & Page 1998), but others are treated as commercial secrets and unknown to the public. Apart from the search engines themselves as a source for how its algorithms work, we therefore have to lean on an important actor in web search, namely search engine optimisation and marketing (SEO/SEM) companies who help web site providers with tactics to end high up in the ranking on a results page. The SEO/SEM sector follows every thread they find on how web search engines rank their results and the changes in these rankings. Search engine companies also directly address the sector. Google, for instance, makes available information specifically aimed at providers of search engine optimisation and marketing services, such as Webmaster Guidelines, a dedicated YouTube channel and other expert advice. At the same time, Google tries to distance itself from the sector in a move to protect their rankings from being manipulated and to project the image of being incorruptible. Google is thus careful to point out that it is not possible to buy a better ranking for a website, at least not an organic one. Google explains how their algorithms for ranking search results work by stating the general criteria the engine applies:

> These algorithms analyze hundreds of different factors to try to surface the best information the web can offer, from the freshness of the content, to the number of times your search terms appear and whether the page has a good user experience. In order to assess trustworthiness and authority on its subject matter, we look for sites that many users seem to value for similar queries. If other prominent websites on the subject link to the page, that's a good sign the information is high quality.
>
> *Google n.d. b*

Thus, Google itself emphasises content, when the content is updated, and popularity in terms both of how often other searchers have accessed a website and how other websites link to a page. One source of information on how factors other than PageRank work are SEO industry publications that often provide information about factors certain to affect the ranking as well as factors they know, and sometimes guess, affect ranking. The factors are often said to be about 200 in number. We will use one company's presentation and roughly summarise these more or less proven factors in the following categories (Dean 2018):

Referring domains
Organic click-through-rate
Domain authority
Mobile usability
Dwell time
Total number of backlinks
Content quality
On-page SEO

On-page SEO refers to the creation of websites in adherence to certain structural and content principles for search engines and people to appreciate them (ibid.).

One of the most important documents in the field of SEO is the so-called *Periodic Table of SEO Success Factors* (Lewandowski 2018, p. 176) hinting at the science-inspired methods the industry has adopted to interact with web search engines. Together with the metaphor of the organic search results this paints a picture of search engines as natural phenomena to be described and tamed with science. The fact that Google names its major algorithm changes after animals (e.g. Panda in 2011, Penguin in 2012; Hummingbird in 2013, Pidgeon in 2014) further promotes this imagery. This works in contrast to the image of the search engine as a library (Lindsköld 2016), which is in many ways the archetype of control and expertise, far removed from the natural wildness of animals, algorithmic or not. This constantly replicated tension between natural and controlled, organic and organised, wilderness and infrastructure is productive in how search engines and in particular Google shape the cultural narrative of their own role.

Algorithms and relevance

The growth of the web is intrinsically bound up with the growing importance of search engines. It is in many respects pointless trying to come up with a figure that quantifies the size of the Web. However, it is safe to say that the growth of the web to the unfathomable size it has today is intimately bound up with the existence of full-text search engines, their increasing sophistication and resulting infrastructural invisibility. Yet, despite this size, today's search engines return their results in a matter of milliseconds. The results page is in fact a composition of results from different databases. Choices have to be made on how to arrange those in relation to each other and responsive to the device used. These decisions are equally prescribed in algorithms. Interestingly, Google also always highlights its speed on its interface. High up on a typical Google search results page search one is presented together with two numbers, one very large and one very small, the number of results retrieved and the time it took to retrieve them. These figures are in most cases useless for the searcher and impossible to verify. Even an apparently meaningless sequence of characters entered as a search term often generates thousands of results and is retrieved in under a second. Popular English language search terms will result in a figure in the hundreds of millions. But what does that even mean in the context of a real search? Can we scroll down and see which result Google decided to rank lowest? An interesting question helping to gauge search engine bias and their conception of relevance would be "Which is the least relevant result of a search according to a specific search engine?" "What is the worst result for a search according to Google?" But, can we even get an answer to this question?

In Chapter One, we introduced the complex and sometimes polarised discussion about relevance and its different understandings that have been going on in information science since the origin of the discipline (see also Nolin 2009). We first

established two opposing views of approaching relevance, subjective (or cognitive) relevance, and system relevance. The former describes what people experience as relevant, while the latter, in its simplest form, designates the relation between a search query, the index and algorithms. Thereafter, we discussed subjective relevance in relation to topical relevance – that is, the de facto relevance of a certain document understood as epistemic knowledge according to the best understanding available. Moreover, we considered briefly the limited interest the relevance literature in information science shows for social aspects. We have now come to a point where we need to add further nuance to this discussion and also complicate it somewhat. After all, what is at stake in ordering algorithms of search engines are decisions pertaining to relevance in one way or another. "Relevance is 'the invisible hand' that governs these systems" writes Tefko Saracevic (2016, p. 1), referring to web search engines and commercial recommender systems amongst others. He distinguishes between five major manifestations of relevance: system relevance, topical relevance, cognitive relevance, situational relevance and affective relevance (Saracevic 2007, p. 1931; see also 2016, pp. 21–22). System relevance also called algorithmic relevance concerns relevance as understood as the relation between query and information in a database; how an algorithm interprets a query in relation to index, data, metadata, and so forth. Topical relevance, even referred to as subject relevance, refers to the relation between queries and a topic/subject. This is close to what Hjørland (2010) asks for, that is what experts in a field socially negotiate and regard as relevant. Cognitive relevance also known as pertinence refers to the "[r]elation between the cognitive state of knowledge of a user, and information or information objects" (Saracevic 2007, p. 1931). Situational relevance, also called utility, highlights the need to see the demands of a certain task or situation in order to assess relevance. Finally and particularly interesting in relation to web-based search engines, affective relevance concerns intents, motivations and emotions regarding information in relation to experienced relevance of information. Saracevic (2016, p. 21) points out that "general social and cultural factors" could be included in situational relevance, but we can also see how these factors are significant for topical or subject relevance. Information retrieval research has investigated and experimented with relevance for decades and continues to do so, but when web search engines came about, much of this research started to be carried out by search engine companies themselves.

To some extent, the relevance that search engines calculate to order results takes all these five aspects into consideration. Google, for instance, states that "[o]ur goal is always to provide you with the most useful and relevant information" (Google n.d. a). On the same page, they explain "[a]ny changes we make to Search are always to improve the usefulness of results you see". The text continues:

> Our engineers have many ideas for ways to make your results more useful. But we don't go on a hunch or an expert opinion. We rely on extensive user testing and have a rigorous evaluation process to analyze metrics and decide

> whether to implement a proposed change. In 2016, we ran over 150,000 experiments, with trained external Search Evaluators and live user tests, resulting in more than 1,600 improvements to Search.

Google refers to "search rater quality tests" (130,336), "side-by-side experiments" (18,015), "live-traffic experiments" (9,800) and "launches" (1,653). The numbers in parentheses are the numbers of tests Google claims to have conducted in 2017. In search rater quality tests, humans provide feedback on how well they experience the relevance of the results from certain searches. In side-by-side tests, different rankings of the same query are provided so as to evaluate the outcome of two different algorithms. In the live experiments, real users are exposed to changes in the results. During launches suggestions of changes in the algorithm are discussed "by our most experienced engineers and search analysts", as Google states it. It becomes clear that, in Google's understanding, relevance is primarily up to the user to decide – what in Saracevic's terms would be cognitive relevance or pertinence – even if they try to model ideal, representative users. It is also clear that the manifestation of relevance as system or algorithmic relevance concerning the application of different algorithms to the index and surrounding data is central to the web search engine. In fact, as Bilić (2016. p. 2) phrases it, "[i]n order to maintain the relevance of its search results, Google needs to promote an ideology of a neutral and objective search engine based on technical innovations". At the same time, Google also tries to make the search experience situational by considering the cultural context, geographical location, language of the user, as well as – to a degree – the search history of the user. It is a way for Google to expand to new countries and areas in the world. This is often referred to as personalisation of search, as we discuss above. The affective aspect of search is more difficult to measure, but we know that users spend more or less time evaluating the search results depending on their motivation (e.g. Pan et al. 2007). More unlikely information seems to get more attention. For instance, in the case of searching for health information it has been found that so-called "query escalation" occurs (Singh & Brown 2016). This refers to situations when an original search for a mostly harmless and common condition is expanded on after the initial results and then transformed into a search for serious diseases, for instance from headache to brain tumour (ibid.). Interestingly, Microsoft holds a patent on a method for predicting query escalation, which is amongst other things based on extracting "escalation features" from web pages and user session histories (Horvitz & White 2015). Finally, topical or subject relevance is to some extent considered through the PageRank algorithm. By letting popularity – that is, the number of in-links and the number of in-links to the first in-links and so forth – be an important factor in ranking, there is – depending on one's understanding of knowledge – some kind of aggregated and collective notion of relevance built in the system. In Chapter One, we referred to how bibliometrics is one of the foundations of the PageRank algorithm, and it is obvious how closely PageRank and citation indexes are related.

If we, as Hjørland (2010) suggests, regard topical or subject relevance as also bringing in a "social paradigm", there are interesting examples of how Google tries to consider

the interest of society at large. These actions seem to follow a pattern, of *public objection* followed by *corporate resistance*, as Mulligan and Griffin (2018, p. 558) point out, in relation to human rights atrocities and specifically the Holocaust. Different ideas of what search engines are and do run into each other. The public perception of search engines as in addition to providing relevant information also acting as "stewards of authoritative historical truth" collides with a commitment on the part of search engine provider to an "engineering logics which tether search engine performance to observational measures of user satisfaction, coupled with limited recognition of the role search engines play in constructing the need being satisfied", as Mulligan and Griffin (ibid.) note.

There are examples when Google more actively tweaks its algorithms in order to give less attention to for instance racist or sexist content. This is not least obvious in the autocomplete and related searches functions, where users are provided with suggestions for how a search query could be developed based on their own search history and other user's data and where certain phrases are in fact blacklisted by Google (Miller & Record 2017). For certain queries they are entirely "turned off". This blacklisting of ostensibly sexist, racist, and in other way offensive phrases could be seen as a way of supplementing the construction of autocomplete phrases based on pure popularity with a societal interest in not exposing users to certain links or, more pragmatically, with a view to protecting their brand. In their Owl project (Sullivan 2017), launched during 2017, Google started to ask users to provide feedback if inappropriate content had been suggested to them in the autocomplete function or in featured snippets (a summary from the first link directly at the search engine result page). Google has also started to assign higher priority to so-called "authoritative" pages, by, for example, asking search relevance raters in their evaluations to be more observant of content that might be regarded as inappropriate. It should not be forgotten that Google's societal responsibility, as manifest with the above examples, is primarily a result of the capitalist logic that underlies its workings. Google earns money on advertisers and offers a free service for content providers and searchers. Anything that sticks out too much, which generates negative media coverage that could engender dissatisfaction among its users is simply bad for business. Accordingly, the ads are also located on the search engine result page with relevance for the user in mind. However, the fundamental conflict between different logics and ideas of what the purpose of a search engine is, between societal interest and quantifiable user satisfaction, is not resolved in this way. The pieces in the relevance puzzle do not always fit together and the dissonance this creates can be thought of in terms of a friction, a notion we return to in Chapter Four. Rather, as it constitutes a form of breakdown that makes the collision between these logics acutely visible, it serves to amplify it (see also Mulligan & Griffin 2018).

Chapter conclusion: Search and sociomaterial information practices

Search engines have the near-magic characteristic of being something else and the same each time they are being used. We can describe information infrastructures as relational, transparent, and constitutive of action. Thus, information infrastructures do not just exist. They materialise – i.e. they are in a process of becoming – and this

happens in social practices and along temporal lines. When analysing how people search on the web, there is a need at the same time to analyse how this searching – and the talk about this searching – are not just shaped by the technology involved, but how searching is, in fact, sociomaterial. However, index, data, and algorithms are for the most part invisible in themselves, but we get to know them from their effects.

Search engines are co-constructed every time they are used through the data that we, as users, are consciously and unconsciously providing them with. When we search for something, the performativity of the search engine constitutes us as searchers at the same time as, with our data, we co-produce the search engine. Our actual doings co-construct the performativity of the search engine. Star and Ruhleder thus suggest that, rather than asking what is an infrastructure, the question "when is an infrastructure?" should guide the study of infrastructure (2015[1995], p. 379) in a way that captures its malleability in relation to practices, situations, and expectations. What does this perspective mean for search and online search engines as information infrastructures? It becomes obvious how search and online search engines cannot meaningfully be thought of as distinct. Search is dissolved into its technical structures. These in turn are actualised when they are used in a way that goes beyond co-shaping (see also Orlikowski 2007). Search engines are constituted in and through use. When they are used, their algorithms "learn" and the data they draw on to deliver search results grows and is refined for the individual searcher through personalisation, but also for the collective and future searches in general. Vice-versa social practices, including, of course, human actors, are (re)constituted through search technology. When searching online for information, we make search engines at the same time as search makes us. For instance, when many people, as Noble (2018) so convincingly shows, repeatedly use racist or sexist language when using Google Search, the search engine feeds this back to its users through its autocomplete function and prioritising of results, thus quite visibly implicating them in strategies of oppression. In other instances, the search engine might broaden the view by suggesting related searches countering a trusted narrative, thus potentially challenging discriminatory practices. Then again, the mere possibility of online searching influences parts of social relations and practices that, at first glance, might seem unrelated to the use of search engines. The point is that the constitutive entanglement of users and search engines occurs in specific situations and in concrete terms, in relation to particular sociomaterial practices and fields of power. The next chapter will tease out some of what this implies in order to discuss examples of how this pervasive entwinement plays out in everyday life and its practices.

References

Altaweel, I., Good, N. & Hoofnagle, C. (2015). "Web privacy census." *Technology Science*. 15 December 2015. Retrieved from http://techscience.org/a/2015121502.

Andersson, C. (2017). "'Google is not fun': An investigation of how Swedish teenagers frame online searching." *Journal of Documentation*, 73(6), 1244–1260.

Ausiello, G. (2013). "Algorithms: An historical perspective." In G. Ausiello & R. Petreschi (Eds), *The power of algorithms* (pp. 3–26). Berlin, Heidelberg: Springer.

Bergman, M. K. (2001). "White paper: The deep web: Surfacing hidden value." *Journal of electronic publishing*, 7(1). Retrieved from http://dx.doi.org/10.3998/3336451.0007.104.

Bilić, P. (2016). "Search algorithms, hidden labour and information control." *Big Data & Society*, 3(1), 1–9.

Bowker, G. C. (1994). *Science on the run: Information management and industrial geophysics at Schlumberger, 1920–1940*. Cambridge, MA.: MIT Press.

Bowker, G. C. (1996). "The history of information infrastructures: The case of the international classification of diseases." *Information Processing & Management*, 32(1), 49–61.

Bowker, G. C. (2005). *Memory practices in the sciences*. Cambridge, MA.: MIT Press.

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2009). "Toward information infrastructure studies: Ways of knowing in a networked environment." In *International handbook of internet research* (pp. 97–117). Dordrecht: Springer.

Bowker, G. C. & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, London: MIT Press.

Brin, S., & Page, L. (1998). "The anatomy of a large-scale hypertextual web search engine." *Computer networks and ISDN systems*, 30(1–7), 107–117.

Bucher, T. (2018). *If…then: Algorithmic power and politics*. New York, NY: Oxford University Press.

Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., & Schnapp, J. (2012). *Digital_humanities*. Cambridge, MA.: MIT Press.

Cormen, T. H., Leiserson, C. H., Rivest, R. L. & Stein, C. (2009). *Introduction to algorithms* (3. ed.). Cambridge, MA.: MIT Press.

Council Directive (2016/1148/EU) of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. *Official Journal of the European Union*, L194, 1–30.

Dean, B. (2018). "Google's 200 ranking factors: The complete list (2018)." Retrieved from https://backlinko.com/google-ranking-factors.

Dutton, W. H., Reisdorf, B., Dubois, E., & Blank, G. (2017). "Search and politics: The uses and impacts of search in Britain, France, Germany, Italy, Poland, Spain, and the United States." Quello Center Working Paper No. 5–1-17. Retrieved from https://ssrn.com/abstract=2960697.

Edwards, P. N. (2003). "Infrastructure and modernity: Force, time, and social organization in the history of sociotechnical systems." In T. J. Misa, P. Brey, & A. Feenberg (Eds), *Modernity and technology* (pp. 185–225). Cambridge, MA.: MIT Press.

Edwards, P. N., Bowker, G. C., Jackson, S. J., & Williams, R. (2009). "Introduction: An agenda for infrastructure studies." *Journal of the Association for Information Systems*, 10(5), 364–374.

European Commission (2018). "Critical infrastructure." Migration and Home Affairs. Retrieved from https://ec.europa.eu/home-affairs/what-we-do/policies/crisis-and-terrorism/critical-infrastructure_en.

Ferragina, P. & Venturini, R. (2013). "Web Search." In G. Ausiello & R. Petreschi (Eds), *The power of algorithms: Inspiration and examples in everyday life* (pp. 107–137). Berlin, Heidelberg: Springer.

Foster, V., Butterfield, W.Chen, C., & Pushak, N. (2009). *Building bridges: China's growing role as infrastructure financier for sub-saharan Africa*. Washington: The World Bank.

Gillespie, T. (2010). "The politics of 'platforms'." *New media & society*, 12(3), 347–364.

Gitelman, L. & Jackson, V. (2013). "Introduction." In L. Gitelman (Ed.), *Raw data is an oxymoron* (pp. 1–15). Cambridge, MA.: MIT Press.

Google (n.d. a). "Focus on the user." Retrieved from https://www.google.com/intl/en_uk/search/howsearchworks/mission/web-users/.

Google (n.d. b). "How Search algorithms work." Retrieved from https://www.google.com/intl/ALL/search/howsearchworks/algorithms/.

Google. (n.d. c). "How Search organizes information." Retrieved from https://www.google.com/intl/ALL/search/howsearchworks/crawling-indexing/.

Google (n.d. d). "Your data." Retrieved from https://privacy.google.com/intl/en-GB/your-data.html.

Hagiu, A. & Wright, J. (2015). "Multi-sided platforms." *International Journal of Industrial Organization*, 43, 162–174.

Haider, J. (2016). "The structuring of information through search: Sorting waste with Google." *Aslib Journal of Information Management*, 68(4), 390–406.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). "Measuring personalization of web search." In *Proceedings of the 22nd international conference on World Wide Web* (pp. 527–538). ACM.

Hjørland, B. (2010). "The foundation of the concept of relevance." *Journal of the American Society for Information Science and Technology*, 61(2), 217–237.

Horvitz, E. J. & White, R. W. (2015). "Predicting escalation events during information searching and browsing." Microsoft Technology Licensing LLC. U.S. Patent 9,015,081.

Keane, M. (2017). "Foreword." In M. Kent, K. Ellis & J. Xu (Eds), *Chinese social media: Social, cultural, and political implications*. New York, NY: Routledge.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Los Angeles: Sage.

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). "Location, location, location: The impact of geolocation on web search personalization." In *Proceedings of the 2015 Internet Measurement Conference* (pp. 121–127). ACM.

Kling, R. (2000). "Learning about information technologies and social change: The contribution of social informatics." *The Information Society*, 16(3), 217–232.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lewandowski, D. (2013). "Suchmaschinenindices." In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen 3: Suchmaschinen zwischen Technik und Gesellschaft* (pp. 143–161). Berlin: AKA Verlag.

Lewandowski, D. (2014). "Why we need an independent index of the web." In R. König & M. Rasch (Eds), *Society of the Query Reader* (pp. 50–58). Amsterdam: INC.

Lewandowski, D. (2018). *Suchmaschinen verstehen* (2nd ed.). Berlin, Heidelberg: Springer.

Lindsköld, L. (2016). "Google as a library: Conceptualisations of Google search regarding the 'Right to be Forgotten' ruling." (Short paper) *COLiS 9 Conference*, *Uppsala, 27–29 June 2016*.

Lindsköld, L. (2018). "Google as a political subject: The right to be forgotten debate 2014–2016." *Online Information Review*, 42(6), 768–783.

Madhavan, J., Ko, D., Kot, L., Ganapathy, V., Rasmussen, A., & Halevy, A. (2008). "Google's deep web crawl." *Proceedings of the VLDB Endowment Archive*, 1(2), 1241–1252.

Mager, A. (2012). "Algorithmic ideology." *Information, Communication & Society*, 15(5), 769–787.

Mattern, S. (2014). "Library as infrastructure." *Places Journal*. Retrieved from https://doi.org/10.22269/140609.

Miller, B. & Record, I. (2017). "Responsible epistemic technologies: A social-epistemological analysis of autocompleted web search." *New Media & Society*, 19(12), 1945–1963.

Mulligan, D. K. & Griffin, D. S. (2018). "Rescripting search to respect the right to truth." *Georgetown Law Technology Review*, 2(2), 557–584.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Nolin, J. (2009). "'Relevance' as a boundary concept: Reconsidering early information retrieval." *Journal of Documentation*, 65(5), 745–767.

Olson, H. A. (2001). "The power to name: Representation in library catalogs." *Signs: Journal of Women in Culture and Society*, 26(3), 639–668.

Orlikowski, W. J. (2007). "Sociomaterial practices: Exploring technology at work." *Organization Studies*, 28(9), 1435–1448.

Orlikowski, W. J. & Scott, S. V. (2008). "Sociomateriality: Challenging the separation of technology, work and organization." *The Academy of Management Annals*, 1(2), 433–474.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). "In Google we trust: Users' decisions on rank, position, and relevance." *Journal of Computer-Mediated Communication*, 12(3), 801–823.

Pilerot, O. (2014), "Making design researchers' information sharing visible through material objects." *Journal of the Association for Information Science and Technology*, 65(19), 2006–2016.

Plantin, J. C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). "Infrastructure studies meet platform studies in the age of Google and Facebook." *New Media and Society*, 20(1), 293–310.

Rieder, B., & Sire, G. (2014). "Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web." *New Media and Society*, 16(2), 195–211.

Sandvig, C. (2013). "The internet as infrastructure." In W. H. Dutton (Ed.), *The Oxford Handbook of Internet Studies* (pp. 86–106). Oxford: Oxford University Press.

Saracevic, T. (2016). "The notion of relevance in information science: Everybody knows what relevance is. But, what is it?" *Synthesis lectures on information concepts, retrieval, and services*. San Rafael, Calif.: Morgan & Claypool Publishers.

Saracevic, T. (2007). "Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and Manifestations of Relevance." *Journal of the American Society for Information Science and Technology*, 58(13), 1915–1933.

Singh, K. & Brown, R. J. (2016). "From headache to tumour: An examination of health anxiety, health-related Internet use and 'query escalation'." *Journal of health psychology*, 21 (9), 2008–2020.

Spärck Jones, K. (1972). "A statistical interpretation of term specificity and its application in retrieval." *Journal of Documentation*, 28(1), 11–21.

Stadler, F. & Mayer, C. (2009). "The second index: Search engines, personalization and surveillance." In K. Becker & F. Stadler (Eds.), *Deep search: The politics of search beyond Google* (pp. 98–115). Innsbruck: Studienverlag.

Star, S. L. & Bowker, G. C. (2010). "How to infrastructure." In L. A. Lievrouw & S. Livingstone (Eds), *Handbook of new media: social shaping and social consequences of ICTs* (pp. 230–245). London: Sage.

Star, S. L. & Ruhleder, K. (2015[1996]). "Steps toward an ecology of infrastructure: Design and access for large information spaces." *Information systems research*, 7(1), 111–134.

Sullivan, D. (2017). "Google's 'Project Owl': A three-pronged attack on fake news & problematic content." *Search Engine Land*. Retrieved from https://searchengineland.com/googles-project-owl-attack-fake-news-273700.

Sundin, O. (2015). "Invisible search: Information literacy in the Swedish curriculum for compulsory schools. "*Nordic Journal of Digital Literacy*, 10(04), 193–209.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices." *Journal of Documentation*, 72(6), 990–1007.

UNASUR (n.d.). "China will increase its economic presence in UNASUR." Retrieved from https://www.unasursg.org/en/node/678.

Zimmer, M. (2008). "The gaze of the perfect search engine: Google as an infrastructure of dataveillance. " In A. Spink & M. Zimmer (Eds), *Web search: Multidisiplinary perspectives* (pp. 77–99). Berlin, Heidelberg: Springer.

4

SEARCH IN EVERYDAY LIFE

In the spring of 2016, a young Chinese man died as a result of a rare and aggressive form of cancer. He had been diagnosed a couple of years earlier and had received medical treatment. However, the treatment failed and he did what most people would do today. He looked online for other options. He searched. To do this he turned to Baidu, the dominant search engine in China. The search engine presented him with results that led him to turn to a treatment centre offering an experimental treatment for cancer, supposedly carried out in cooperation with Stanford Medical School. This cooperation turned out to be pure fiction and the treatment, while very costly, was unsuccessful. The young man passed away. Before his death, however, he posted a long message on *Zhihu*, a question and answer website telling his story (*China Daily* 2016). In short, he accused Baidu of corruption, of manufacturing search results with commercial interests in mind. The search engine failed to label adverts clearly amongst search results, and thus made it nearly impossible to distinguish paid content from other results. They appeared to the searcher as part of the organic results. This led to a social media storm, complete with a dedicated hashtag. The case was picked up by national and international news media as well as by the Chinese authorities. Apparently, there was in fact a problem with how Baidu's algorithms produced the order of links on their search engine results page. It seemed that advertising money could indeed buy ranking order and the suppression of negative reviews.

Already before this tragic event, Baidu's role in how information on health related issues is searchable (or not) on Baidu has been a much discussed issue of contention. Browsing through the pages of the English version of *China Daily* – which does convey the Chinese government's official view – shows that problems with the search engine's role in everyday life, a supposed lack of information control or dubious business practices, are often presented in relation to health information. Specifically, the way in which private hospitals and Baidu cooperate

DOI: 10.4324/9780429448546-4

and Baidu's control over how health care providers can be searched for online is mentioned repeatedly. It is thus no surprise that Baidu's malpractices are discussed in relation to health information on *China Daily*, and that the call for stricter rules and increased control is articulated by reference to information on health and health care. This is couched in a terminology of objectivity versus corruption, as also purported by the "Cyberspace Administration of China" (CAC). A headline, echoing a report by the CAC, reads for instance: "Baidu's objectivity compromised by profit model: investigation" (*China Daily* 2016). This raises the bar very high for what a search engine should deliver and points in an interesting direction. Clearly, objectivity in this example works mostly as a discursive signal for technical neutrality and non-commercial interests rather than objectivity in a more philosophical sense, but still it signals that the ideal for information technology is to be a neutral broker of information with no external factors corrupting it, especially not money. Legal frameworks and regulation and even direct intervention through the state, as is the case in China, in this reasoning appear thus less as an active shaping of information or information control and more as a guardian of neutrality, or even objectivity.

Was this what the young man in search of a cure for his cancer was looking for; neutral information ranked in ordered lists only determined by objective factors? We have no way of knowing, but it is likely. As we saw in Chapter Two, most people seem to have little understanding of how search engines order their results and in which ways they embody value judgments. At the same time, one of the reasons why search engines have become engrained in all kinds of everyday life practices is precisely because they are not "objective", in any meaningful sense, but increasingly successful at second-guessing user needs without users even articulating those needs in the first place. That is, even if a search query is entered, users neither actively instruct the search engine how to interpret this query nor how to order and return the results. The way search engines interpret users' needs could be related to what has been driving librarians for decades, to be able to provide users with literature or information that they "need" before they realise the need themselves.

In a classic article from 1968, Robert S. Taylor (1968) visualised how information needs develop from something that is largely visceral, rather than intellectual, into something quite specific, which it is possible to express in a way that is meaningful for an information system. Yet, as we touched upon in Chapter Two, what actually constitutes a *need* is not easily determined. According to Taylor, a need develops from a gut feeling and turns into something mentally explicit, which in the end can be expressed in words that can, if required, be entered into an information system. However, there are also other ways of referring to needs in relation to search engines. When weighing up the crucial significance of the order of search engine results, we might ask in what ways search engines also encapsulate a societal interest. As we mentioned in Chapter Three, Google prioritises to some extent links to more established, authoritative websites, since a measure of authority in this context includes for instance the age of a domain or the consistency of its content (Lewandowski 2018, p. 180). Yet, for most intents and purposes, the search

engine does not care about society as long as the search result satisfies a "need", that is, something the individual user experiences as relevant (or pertinent). Occasionally and typically accompanied by public outrage, the media expose examples of conflicting interests of individual needs and societal interests. Typically, this is connected to exposure of search results that are perceived as extreme or which are manifestly wrong, but we rarely think of them in everyday life. As we discussed in Chapter Three, these extremely obvious examples make the infrastructure and its embedded algorithmic decision visible. They are experienced as forms of *infrastructural breakdown*, articulated and palpable as, what we suggest calling *frictions of relevance*. The notion of frictions of relevance in the case of general purpose search engines describes the dissonance of different individual needs, societal interests, and the vested interests of the stakeholders catered to by the business model of the multisided platform structure.

From specialised to mundane

Can searching for information ever be regarded as purposeless or be initiated without a trigger? The literature on information seeking in everyday life (e.g. McKenzie 2003; Savolainen 2006) asks us to investigate information seeking beyond task-initiated activities in working life with a starting point in the rational mind of an individual. Watching TV after dinner, listening to the radio while driving to or from work, reading the daily paper over breakfast, and similar activities could easily be understood as being without explicit purpose beyond being part of established social practices where they are one of a series of stable elements. They are predominantly grounded in routines rather than carried out in response to an explicit trigger such as a task, a problem, or an experienced gap. Andrew Cox spells out the difference between a practice approach and traditional information behaviour research in the following way:

> An obstacle to the adoption of practice ideas is the way that the information behaviour area of study has taken as its starting point individualistic notions such as information needs, and that these are so deeply ingrained into this approach that it is hard to incorporate ideas with such fundamentally different assumptions.
>
> *Cox 2012, p. 184*

A practice approach takes its starting point in "what people normally do" rather than in individuals' information needs. As we discussed in Chapter One, social practices are generally theorised in terms of recurring patterns, and a practice is thus perceived of as being constituted of a consistent and ordered nexus of activities, artefacts, people and, woven across it – information. Sanna Talja and James Nyce succinctly formulate what is at stake (2015, p. 65): "[w]hen studying everyday practice, a central concern is how social, material, and technological resources help to keep the practice going and to modify it, and how these offer constraints and possibilities for action".

What happens if we apply a practice approach to the study of online searching? After all, this requires people to (at least so far) interact with an information system, an activity that demands an active start by entering words into an empty white

query box and pressing the search button or speaking a query out load after asking to initiate voice command. At the same time, the situations search is used in are vastly different from each other, and this needs to be accounted for, as the practice approach also advocates. It is thus important to take these differences seriously and thus how people reflect on online searching and how they differentiate between different situations. What do they search for, when and why, and when do they not search online? We posed these and similar questions to people in Sweden in a series of focus groups. The participants, whose ages ranged from 13 upwards, in these conversations on search came from different walks of life, and had different nationalities and professional backgrounds (Andersson 2017a, 2017b; Carlsson & Sundin 2017; Haider 2017; Kjellberg & Haider 2018; Sundin et al. 2017; Sundin & Carlsson 2016). What we found is not just how online searching has come to constitute a central activity of everyday life – a search-ification of everyday life – but also how it has become so ordinary, so nested in our daily practices and so easy to do, that we could talk about a mundane-ification of search (Sundin et al. 2017). What used to be a highly specialised activity, in fact something that could be seen as a practice in its own right, as we touched upon in Chapter One, has for most intents and purposes turned into an activity embedded into the various practices of everyday life. Here, these conversations provide us with snapshots, taken at one moment in time in 2014 and 2015. They help us to understand how online search is embedded into everyday life in rather specific ways and how this has implications for how we deal with information, individually and as a society.

We suggest that online search, in many instances, is so deeply embedded in everyday life that it is near impossible to tease it out as an information practice that can stand on its own. A lot of research in information science in recent years has recognised a similar dilemma (e.g. Lloyd 2010, 2014; Rivano Eckerdal 2012; Talja & Nyce 2015; Veinot 2007; Pilerot 2013). It has then moved its focus from information in itself to social practices and specific settings, in the course burying information-related issues deeper into the fabric of social life and also, as a side-effect, making information technology increasingly invisible. We are not suggesting returning to a prior paradigm where information can be singled out and then investigated as independent of its context and its use; rather, we want to take the practice approach closer to technology, to its control and also to society in a more profound way. It is an attempt at delivering an understanding of the role of online searching in everyday life emanating from a sociomaterial understanding of practices and which is responsive to the agency of non-human actors, such as, for instance, algorithms, smartphones or indexes and also considerate of the corporate uniformity of these platformised technologies. The ease of online searching paired with mobile technology has made the activity of searching for information a potentiality in almost any social practice. What previously, at least to some extent, could be interpreted as an information practice in itself has now often become integrated into other practices. At the same time, it has also kept some of its complexity. Yet, how more exactly does this embedding occur? Consider the following exchange:

PERSON 1: It has sort of become a habit, a routine in some way. It's not like "now we're going to search on the internet".
[Everybody laughs]
PERSON 1: "Let's gather the family!" No the early computer and those CDROM discs and you put it in the computer and you got to see a movie of a lizard running.

The quote illustrates in a tongue-in-cheek and perhaps slightly exaggerated way how earlier technology could make the activity of computer-mediated searching into a goal in itself – to gather the family around an encyclopaedia to watch a lizard running, while the ubiquity and routinisation of search has made it into something else. For the participants in our study, a major change came when search became mobile – when the smart phone became the new companion:

PERSON 1: I think, for me, when I moved to Sweden like a year ago ehm I got my first smartphone and before I didn't really need a smartphone I didn't, you know, if I wanted to find out information I would do it before I went out and now I use my phone for everything, you know.
PERSON 2: Once you have it you can't live without it.
[laughs]

This is a good example of how the possibility of searching online changes how we do things and how this is – also – a profoundly temporal change that has to do with how social practices are conceptualised. That is, *when* a search is carried out makes a difference for the make-up of practices. In our empirical material, what is hinted at are little things, like looking up an address, a map, opening or travel times, taking down notes, printing out, or remembering, and so on. All these are small activities that constitute the structure of a practice, of for instance visiting a new acquaintance or going shopping or whatever else it could be and moving or replacing them has implications for the rhythm of everyday life. Certain encounters with people are replaced with googling. Someone talks, somewhat regretfully, about not speaking to her mother as often as she used to because Google helps her answer certain questions faster than a phone call could have. One participant reflected on how the use of Google changed his relationship with a colleague who he used to ask for technical advice and thus his expectations of how fast a request should result in a satisfying answer more generally: "This generates a certain restlessness in oneself" is the expression he uses. Also here what we are confronted with is a temporal and also an affective change in how social relations play out, which is facilitated by the use of web search engines.

Another participant made an unusual and colourful comparison, also illustrating the power of mobile search: "I remember once when I had my first iPhone 3 and was like this … yes, it was almost like masturbation like I had to daub and look at the screen and look for new things all the time". These inconsequential lookup searches, often made without being noticed at all, or at least – when doing them – without being further reflected upon, seldom lead to any evaluation of the

information found. It does not really matter whether the information is "true" or not, as long as it does what it is supposed to do. Another participant reflects:

> So, because it is so easily accessible and such a natural part of everyday life that you maybe don't reflect or think twice in some situations or with some searches which you perhaps should do.

To reiterate, the search-ification of everyday life has turned a once complex professional practice into something seemingly simple and made it available for everyone. A search can be carried out quickly and easily, and often without any need for time to reflect. The obligation to evaluate information found through a search engine is constructed in relation to the situation one is in and to how consequential something is perceived to be. In the words of a participant in one of our focus groups: "when it's just trivial things you're looking for, then it doesn't really matter how true something is – maybe". In Chapter Two, we showed how evaluation in fact seems to be carried out by people when they consider the question to be important. At the same time, the search-ification of everyday life also means that, when searching forms a part of the mundane, it concurrently becomes more difficult to discern – not just for researchers investigating the phenomenon, but also for all of us who search for information.

When the starting point is in ubiquity and routinisation, the very idea of critical evaluation of information changes. How could you evaluate searching, when search has fallen into the background of everyday life to such an extent it is hardly recognised at all?

PERSON 1: But I think like I said that it's routine, you know, routine over the whole thing. Before it was a bit exciting to use the computer but I mean now it's like I read it like I read the morning paper, I just sit there and scroll.
[Inaudible]
PERSON 2: I haven't really reflected on it, like why does that come up on top or maybe I do sometimes but I think that you just trust Google.
[Laughs]

In general, people tend to trust what Google provides them with, without scrutinising what they get (and do not get). Ralph Schroder (2015, 2018) rightly questioned traditional information retrieval evaluation as a starting point for understanding evaluation of web search, since so much web search is not actually about searching for epistemic content of social importance (Schroeder 2018, p. 125), but rather a search for information related to amusement and leisure. However, as we argue in Chapter Two with reference to Vivienne Waller (2011), these kinds of searches are still carried out in their millions every minute around the globe. Traditionally, information retrieval research is concerned with finding documents of a serious kind. Even if these experiments are carried out in labs, the self-evident understanding is that search is a complex task carried out by specialists in a research or professional environment. On the other hand, when information

retrieval has become a web phenomenon, so-called socially important searches could also be understood to be routinised to such an extent it is almost invisible. In a way, web search engines in everyday life tend to link "serious" and "leisure" so that the two categories are blurred, leaving evaluation to the ranking algorithm of the search engine. We attend to the practice of critical evaluation of information in more detail in Chapter Five.

The ubiquity of search

While it is interesting in certain respects that web searches for content related to e-commerce and popular culture vastly outnumber web searches for other content, in everyday life it is difficult to separate the different categories from each other and it is not always meaningful to do so. This is in fact one of the characteristics of everyday life – its messiness. Facts can be entertaining, and popular culture can for sure be facts. The routines of family life, the changing of the seasons, the coming and going of holidays, the course of life, as grand as it sounds, from birth control to bringing up children to aging, and even death and mourning, all these are implicated in how technology is used, and search engines are no exception. Most people do not normally search for cures to life-threatening diseases, as in the example at the beginning of this chapter, but it happens and when it does, it can have consequences that result in life or death. The distinction between lookup search and explorative search, as provided by Marchionini (2006), could be used here. People search mostly for answers to short lookup questions, but we also make explorative searches involving more complex issues, while a great number of searches can be placed in between these types. Logfiles, Google trends, and various statistics all paint a certain picture. Online search is everywhere and always, so much so that search engines are now used by the media as cue givers for interviews; the number of searches for certain conditions is used in some countries to fathom whether seasonal diseases are underway (Carneiro & Mylonakis 2009), attempts are made at using Google to predict election outcomes (Mavragani & Tsagarakis 2016), or love interests are checked up on with a search (Gibbs et al. 2011), and school children have begun to find Google boring (Andersson 2017b). People's experience of the internet is measured and detailed regularly by research institutes around the globe and presented in reports and subjected to statistical analysis. Interestingly though, while the number of such reports dedicated to social media is enormous, those that investigate search engines and their use are far and few between, and even in general "internet use surveys" search engines are not often mentioned beyond an almost duty-bound mention of Google's market share, sometimes followed by mentioning also-runs Bing and Baidu. The World Internet Project International Report, issued each year by the Centre for the Digital Future at USC Annenberg School for Communication and Journalism in the USA, bringing together data from very different countries, while an interesting source in many other regards, is an example of this subtle submerging of search engines into the

infrastructural background. One exception is the 2012 report by the US American Pew Research Center, which summarises its major findings as follows:

> For more than a decade, Pew Internet data has consistently shown that search engine use is one of the most popular online activities, rivaled only by email as an internet pursuit. In January 2002, 52% of *all Americans* used search engines. In February 2012 that figure grew to 73% of *all Americans*. On any given day in early 2012, more than half of adults using the internet use a search engine (59%). That is double the 30% of internet users who were using search engines on a typical day in 2004. And people's frequency of using search engines has jumped dramatically.
>
> *Purcell, Brenner, & Ranie 2012, p. 3*

The same report shows that 91% of respondents think they find what they are looking for most of the time. When writing this book, there is no new report from Pew Research Center on search engine use, so we cannot follow up the figures from 2012, and increasingly search engine use seems to be implied in similar surveys when measuring, for instance, the use of "the web" or "the internet" for fact checking, for looking up information and so on. This is indicative, we suggest, of at least three issues; firstly, the increasing mundane-ification and ensuing simplification of search and search engines, secondly – at the level of survey design – the difficulty in measuring, talking about, or even just perceiving of an activity that is hardly ever carried out as an independent one, but almost always in relation to something that makes it change its shape, and thirdly the problem in pinpointing an activity that, in large parts of the world, is so closely tied to just one brand name that all questions about online search in general as opposed to googling have an air of forcedness about them, making them appear close to meaningless. However, we could guess there has been some competition between search engine use in the traditional sense and the development of mobile phone applications or smart speakers that further submerge search engines to a degree that makes awareness of them increasingly difficult. Also, the growing popularity of YouTube might have an effect on search engine use (Ofcom 2018). Furthermore, predictive automated search is slowly emerging as a possibility, where searches are carried out in response to questions the search engine poses itself on behalf of the user without the user even being aware of this happening – here the agency of non-human actors and the constitutive entanglement of the social and the material which sociomaterial theory describes (Orlikowski 2007) becomes acutely obvious. It is until now often related to travel time and couched between the calendar application and the map application – conveniently part of the same platform and typically installed on the same phone. A more recent, example where search engine use actually is surveyed and mapped in quite some detail, is a report issued by the United Kingdom's regulator for communications services, Ofcom (UK). This report (Ofcom 2018) shows that the popularity of search engines is very high: 95% of the respondents report they have used general-purpose search engines, which makes it by far the most popular source for information online.

So, what could search in everyday life actually be? As a trigger for the conversation, each of our focus groups started by the moderator asking participants to write down different recent occasions when they searched for information. Below is an example of a typical search note:

1) Title of a song
   Google voice searched first lines of song. Behind blue eyes.
2) Dinner inspirations
   Searched on instagram #dinner.
   Searched on Google what to cook for dinner
3) Bus info
   #swebus on Twitter see how it looks like, other info on Google.
4) [name of condition].
   Searched symptom.

This illustrates quite well how the variety of reasons for searching and popular culture can be mixed with information on medical conditions, and thus elucidates what we call the search-ification of everyday life. A song title, dinner inspirations, bus information, and symptoms of an illness – almost anything can be searched for. Or, as one of the participants answered when asked when he did not search, "[w]hen the internet is down"; another one postulated "[w]hen the data allowance is used up" in reference to the person's mobile phone contract. The note also shows that it is not only Google that mediates search but in this case also social media, such as Instagram and Twitter. Search is, of course, not equivalent to googling, but Google seems to colonise other ways of finding information in everyday life. Google is also used to search on other websites, even when it is known there is a site for a business and there is a search mechanism on this site. This type of navigational search is for instance at issue in the following quote: "What has changed most in my searches is that I can't cope with going to the web pages directly, but I just write in what I am looking for exactly", referring to Google. Google also transgresses the boundaries of different practices, as this lawyer phrases it in one of our focus groups:

> That has changed very much in recent years /…/ all the time /…/ I believe you don't even notice how much more you find nowadays. I remember when I worked at the district court in 2004 and 2005 when it was still kind of not new, but a bit unusual with Google and it was sort of a bit sensational when a judge actually googled, because one was so used to the idea that you should always go to reliable sources and sort of search properly /…/ and now you google countless times every day, even if you of course still use proper, reliable databases for some searches.

The above account is an interesting example of the context collapse that the dominance of general-purpose search engines in general and Google in particular have brought about. It is the same engine, the same interface, the same mechanisms

encountered in both professional situations, highly qualified ones even, as above, and all aspects of private life. The notion that an infrastructure is learned as part of a community of practice, as discussed in Chapter Three, becomes problematic, as in the case of Google it is the same infrastructure interlacing different communities of practice. If, as discussed in Chapter Two, we regard searching as being located within a library paradigm for finding information, social media lies closer to a so-called village paradigm (Trias i Mansill & de la Rosa i Esteva 2013). Another participant:

> For me like when I need recommendations I won't go on Google. I would go on Facebook and ask like what's the best restaurant to go to or something or like when it's quite subjective and ask something that's so open then it's quite hard for me to go to Google because /…/ there are certain determinates why these are the top searches but if I know like there are some people in my circle that yes I would feel that we have the [same] interest and I would trust that more.

Facebook combines, as is well known, a constant flow of information with the possibility to ask directed questions of trusted friends. Where Google's search results are not (yet at least) personalised to such an extent they can challenge the status of known people, Facebook or other social media operationalising the above-mentioned village paradigm step in. Thus, for recommendations such as for restaurants, the authority of friends wins over the authority of Google's algorithm – at least in certain situations. In other situations, this complementary relationship comes to the fore in a different way, and it is increasingly common to insert a "Google it" in a Twitter or Facebook thread to cut short an argument, taking advantage of Google's established status as a neutral arbiter of facts and opinion. Search, not just the use of general-purpose search engines, but also of other online search tools, is embedded into and given meaning in relation to networks of other information-related activities, online and offline, analogue and digital. Increasingly, search engines unearth and prioritise results from user-generated content, like Wikipedia, but also blogs, Q&A forums, and such like. In a certain way, the village paradigm and the library paradigm merge into the mirage of a *village library*, a privately operated one, unaccountable to its constituency, but taking advantage of its members' informational work and most of all, the trust-building work that makes them useful.

The search-ification of everyday life as well as the mundane-ification of search is closely linked to the constant availability of internet access mediated through smart phones, which has turned searching into a possibility everywhere and anytime:

PERSON 1: Yesterday I was online because I was going to cook halibut and then I thought that I would make a hollandaise sauce but I couldn't remember how to make it so I went to recept.nu [a website] and then made it.

PERSON 2: I was online on my phone checking when it's time to crop lavender; it's during the autumn.

Online searching is embedded across different types of practices, carried out in very different places, and at various times; in the kitchen, in the garden, related to the routines of cooking and the rhythm of the season. These examples seem innocent, and in many ways of course they are. Yet, the reasons why a web page ends up at the top of a result page are multifarious and they have consequences. The worst outcome in the above examples might be a spoiled dinner party or a plant that fails to grow, yet, for people looking for medical advice or for those behind the publication of recipes or gardening advice, the consequences could be substantial. The type of quick lookup search that is so common in everyday life, where the smartphone together with Google act as an external memory as in the examples above, often ends with clicking on one of the first links at the result page or increasingly the answer is right there, as a snippet or knowledge graph answer on display. Contemporary commercial search engines are, as we discussed in Chapter Three, multi-sided platforms. As such, their business model is based on catering to different user groups (Rieder & Sire 2014). They convey relationships between these groups and part of this relationship building effectively means locking the groups into a form of mutual dependence.

Not searching, searching differently, and avoiding search

Typically the normality of something, its ordinariness, becomes noticeable most acutely when it is not available or when it is disturbing something else, thus requiring us to consciously reflect on and modify what we normally just do without thinking about it. In the previous chapter, we introduced as the best-known characteristic of an infrastructure, the fact that it becomes visible first at breakdown. If we take seriously our own claim that an information infrastructure comprises its use, its embedding into practices, as much as its technical, material structures, then breakdown, as enabling visibility or noticing, can also be imagined at the level of non-use. Thus, we suggest, the breakdown of search engines as an infrastructure is not limited to their technical failing, but needs to be understood in a considerably broader sense that is responsive to the specific ways in which they are entangled with social practices and their routines. How does that happen? When do we not search or when do we search differently?

As J. David Johnson (2009, p. 601) highlights, "ignorance and information seeking are inextricably intertwined concepts" and in information science research as in Western society more generally an underlying tacit assumption is that more information is always to be preferred and knowing is in principle always better than not knowing. Consequently, in the field, the notion of information avoidance is often brought to the fore as an accepted phenomenon, but one that needs to be overcome, a problem to be tackled by the right changes either in the mode of delivery or in the attitude of the group or person identified as having this impediment. The ideal is the enlightened citizen making rational decisions based on factual knowledge acquired by means of a comprehensive search and judged according to the criteria of critical evaluation of information (see Chapter Five).

However, in many situations, not wanting to know, avoiding information is a perfectly rational decision (Case et al. 2005). The classic example of this would be information about a disease, where not wanting to know and avoiding finding out is sometimes a very reasonable course of action. Consider, for instance, this quote by one of our focus group participants who talks about how she deliberately did not search online for information on a disease that a seriously ill friend is afflicted with, because she "was afraid to get too much information or to get too worried". Here, searching is avoided in order to protect one's emotions (see also Myrick et al. 2016). Our interest here lies not in discussing information avoidance as such, but using the notion as a crutch for elucidating how online search engines are so embedded into our various social practices that a deliberate decision has to be made to remove them, not the other way around. A conscious decision to not want to know, either in order to avoid emotional distress or simply because knowing could spoil an experience that was planned for a later point, frequently goes hand in hand with a strategy to avoid search engines, or to modify their use. One of our participants talks about a situation when she searched for information about a TV series she had followed. She actively avoided looking at the results high up on the results page in Google, fully aware that it most likely would be a snippet from Wikipedia giving away a piece of information – how the season ended – she wanted to avoid learning about at this point in time. While this search was actually undertaken, the person consciously, based on her experience and knowledge of Google's algorithms and interface, adapted her behaviour actively, thus excluding certain content, avoiding information.

Certain social situations are considered unsuited to searching. Someone taking out the phone and looking up the answer and blurting it out might disturb the intent of a conversation. The search engine via the smartphone can be seen to kill conversations, as a participant in one of our focus groups aptly described it. Consider the following story told by another participant:

> and I was hanging out with my friend [name] who one evening when we sat drinking tea started to think aloud about, when did he even get himself a mobile phone and that had been maybe only five years earlier or so, but then we sat and talked and he said: "I have thought about this thing about the water tower, how do they actually work and what are they good for and can the technology be developed further, and do they have to be so damn high?" and I was just like full on and was going to find all the water towers in the world and he just: "No [Name], can't we not just talk about this for once, based on the little knowledge we have about water towers and talk for a long time and just let it be and then perhaps we can find out if what we said was gibberish or if was actually close to the truth, because this feels developing" and this I have adopted in a way.

Some occasions lead to a conscious consideration of the role of online search in them. Some questions are not posed in order to be answered, even factual ones; rather talking around and about them is what is seen as the point of certain encounters. The

emergence and popularity of concepts like "device-free dinners", "no smartphone challenges", and the larger "digital detox", while not exclusively aimed at removing search engines from social situations, are all witness to this understanding of the search engine as a potential conversation killer.

In other situations, search engines are really not available, simply because there is no internet connection, or it might not be search engines in general but Google that is not available, because access to it is blocked, as it is in large parts of mainland China. One of our participants talks of an experience travelling in Cuba without ready access to the internet on her phone, unlike what she was used to:

PERSON 1: I realised during travelling how dependent I am on this quick information source so like sometimes you want to know something and you just google it and then it's fine but when travelling you don't always have the internet and in Cuba for a long time I had no internet at all so I wrote down I had kind of a list of things to google as soon as I [laughter, inaudible] then I don't forget about it cause I really wanted to know.

PERSON 2: And then you googled them when you were home?

PERSON 1: Yes or like in the next hostel where I had a computer and the internet I ran through my list and googled all the things I wanted to know.

PERSON 3: Like how was it here I came across this and that oh you've never thought about the option of let's say taking the ferry instead of that train why not check it out when you have internet again and so on.

PERSON 1: Yes or like you see an old building and you really want to know what's the story behind that or whatever and then you google it later on so I write it down that I don't forget.

Another participant talks about her husband travelling in China and, since Google cannot be accessed there and he was not comfortable with using the Chinese search engine or Bing, he used to telephone her and ask her to do certain searches in Google on his behalf back home in Europe. What unites both accounts is that, rather than replacing the missing search engine with, for instance, exclusively asking people, using the locals or another available search engine or a guide book, being a tourist and travelling are for them so bound up with using search engines or Google, that they use quick fixes to botch the search engine and thus the social practices their use is part of. A piece of paper, the telephone, other – older – media forms are employed as place holders, thus not just repairing the missing search engine, but with it also the routine of being a traveller that they are accustomed to.

The last example of it having to be Google is once again witness to the enormous power the brand and the search engine hold. In many contexts, not using Google specifically disturbs a silent norm. A norm which to break is seen as so striking that doing so even deserves media attention. A Google search using the phrase "Alternatives to Google" results in hundreds of thousands of results, as does the same search in any other general-purpose search engine, and also Google News comes up with links to thousands of media articles in response to such a query. A

particularly striking title adorned a 2018 article in *The Guardian*: "I've tried to dump Google and Facebook. But it's been painful". (Collinson 2018), a short piece by their money editor telling of his experience of not using Google (and Facebook) and describing it in physical terms as literally hurting. This is a somewhat random example, but it helps to communicate how disturbing the dissonance that avoiding Google creates can be. Consider in this light the following exchange in one of our focus groups:

PERSON 1: If I use a different search tool [Note: different from Google] then mainly those that anonymise, like [inaudible] and so on?
PERSON 2: DuckDuckGo?
PERSON 3: It has maybe improved by now, but this is a catastrophically bad search engine.
PERSON 4: But why does one want to be anonymous?
PERSON 1: Because we live in a society, which is incredibly unpleasant.
PERSON 3: Exactly.
PERSON 4: Why do you then want to be on the internet in the first place? It feels like you are surveilled in any case so it doesn't matter which…
PERSON 1: But the internet is fantastic, a fantastic tool. This is like saying why don't you want to read books. Why do I want to read a book… because reading books is fantastic; why don't I want a damn bunch of [inaudible] who stand there and look over my shoulder while I read my book?
PERSON 4: Yes okay, sorry. But there I understand that you want to read a book, but you probably don't want to know, probably don't want to go and get it from a library because then it ends up in a database which book you read, you'd probably rather…
PERSON 1: No it would be good to be able to…
PERSON 4: Smuggle it out from there and…
PERSON 1: Can't I just read my book in peace, is the question I ask…
PERSON 4: And why do you think you are less surveilled if you use a different search engine?
PERSON 1: Because they are dedicated to anonymising and all that. A practice is disturbed and the person doing this on purpose is questioned about it. Not using Google is discussed as a difficult and abnormal situation to be in and if it is a choice, then the person has some explaining to do.

The search interfaces of general purpose search engines are extremely simple to use; you write words into a small query box and hit the search button or, increasingly, speak the search aloud to a voice assistant. However, this does not mean that everything can be searched or that we always know exactly how to formulate the query. One of our participants described a situation where she was confronted with a vegetable she did not know and therefore had no idea how to cook it. She had no idea how to search for it. The interesting point for us here is that she describes the situation as frustrating, since the normal, the obvious thing to

do find out about something – consult a search engine – was not available to her. She was not able to formulate the query and thus had to break a routine. In another conversation a woman puts into words the frustration she experiences when something cannot be searched:

> I thought that you expect that all activities must have a web page, even flea markets and other things should, so you get frustrated when you know that there is a flea market and you just want to know if they are open or not and you don't find it.

We can imagine other situations where this could occur, sometimes the simplest interface, the easiest most straightforward search, fails. An event or suchlike simply does not have an up-to-date webpage or is not included on any listing that can be indexed by the search engine. These situations and the frustration they engender work to magnify how normal and embedded the activity is into everyday life on different levels. To sum up, making visible the void that not searching online creates in a certain situation works to nuance the make-up of a practice and it shows how profoundly search is entangled across it. That searching is much more and something different from merely looking up information understood as an epistemic content is especially tangible in situations where it is either not available or avoided, as is the infrastructural function of search engines in everyday life. The use of online search engines is not necessarily rational, but also affective (Huvila 2016). It is implicated in how we negotiate emotions, relate to ignorance and knowing, or to participation in society and just being sociable. The practice is configured in a way that makes searching online the normal, the expected thing to do or at least the activity requiring least effort in these situations. Disentangling search from everyday life almost leaves a scar; disentangling Google – if we recall the *Guardian* headline mentioned above – definitely seems to do so.

Searching for confirmation: Information feedback loops

The societally significant role of search engines is often discussed in terms of how search engines shape the communication of information and its availability. These discussions, some of which we touched upon in Chapters Two and Three, revolve around the issue of algorithmic power and biases and often concern how search engines prioritise certain types of content, including critical discussion of the personalisation of search results. These reflections offer important insights into the politics of search engines and the various structural biases underpinning online search. However, as we also touched upon in Chapter Two, what these accounts often lack is a consideration of how the very act of carrying out a search is part of the social shaping of search and how this facilitates what could be described as information feedback loops. Consider the following statement by one of our focus group participants:

PERSON 1: It's also if I need, like I need more sources. Even if I've found one good source I still need something to back it up. Then I'll use Google to see if I can find something there that also backs it up.

Here a web search is carried out in order to corroborate something already known (see also Andersson 2017a). In other words, the search engine is asked to provide a confirmation, not support for something the person doubts and wants to know more about, but a source to confirm information that is already accepted as correct, because the social situation, the practices of homework in this case, demands more than one source. Another participant talks very similarly about how she searches online to confirm something known to and accepted by her, but for which she needs help in convincing others. Web search engines help finding support for "what you think you know", as a participants formulates it. Another of our participants expressed this way of searching as follows:

> [T]his is always so interesting that how what you [inaudible] information searching, if you think about it a little bit, that, depending on which search term we use, – you talked about this sort of, how shall one find it, the frustration sometime about it – and the search terms are also something that can be ideological or value-laden or I am going to be like this, I mean, you have to use words, you can't just sit and think, then you find nothing in the end, so you just have to use words I mean, but when you look at it from a meta perspective…

In the same group, another person elaborates on this by describing how she searches for a foodstuff she considers unwholesome: "Then I don't sit there and google 'milk good for your health'. Then I google 'milk bad'". As Andrei Zavadski and Florian Toepfl (2019, p. 33), mentioned in Chapter Two, remind us: "Search results are largely predicated upon how users phrase their queries". In the above example, this characteristic of online searching is used consciously, in a strategic way to support a claim, and the search engine turns into a tool for confirmation bias. In a study on news consumption by conservatives in the USA, Francesca Tripodi (2018, pp. 25–34) made a similar observation. The participants in her study also used search engines, Google more specifically, as tools to confirm opinions and to validate claims. They did this not least by pasting verbatim quotes from sources they already trusted into the query box. The phrases used for carrying out searches have – of course – a very direct impact on which results are returned by the search engine and in which order they appear. Thus, as Tripodi (2018) also highlights, even if people search diligently and try to find different sources in order to validate a claim, if the queries are formulated in accordance with a certain position – political or otherwise determined – then the searches work to a considerable degree to affirm this very position. This is done deliberately, as some of the participants in our focus groups witnessed to, or intuitively as the results in Tripodi's study seem to suggest. At the same time as this query-induced information feedback loop

occurs, choosing search terms is shaped by the affordances that the search engine's interface offers and by people's assumptions or knowledge of how web search engines operate.

Furthermore, search strategies are adapted to generate results that fit the purpose. In a focus group with academic researchers, one participant declared:

> I mean the development in society in general is towards increased efficiency and eh, so there's like no time to follow up on certain threads like for example if I don't find it, ok then I'll chose another way to approach the topic in a way where I might find something online.

Shortly afterwards, in the same focus group, another participant explained, "[it's] not the reflection that carries out the search, but rather you search while reflecting on what is possible to find". Searching, reflecting, and problem formulation are intertwined and in a way they can be said to be each other's outcome. This lends a new twist to the discussion of how search engines are biased and of how they work to advance and solidify certain positions. Choosing search terms to be submitted to Google or another general purpose search engine is a far from disinterested activity. Rather it is – also – part of how opinion is articulated and positions are reinforced. Queries are shaped to fit the search engine's technical affordances, different understandings or sometimes imaginings of how they work, and the various social, cultural, and political arrangements we relate to, and they have a part in our positioning vis-à-vis those arrangements. Conversely, information, as content, is increasingly produced in direct response to search engines' various technical arrangements, either to be found by means of search engines or to be hidden underneath their surface. Taken together, it is hard to overstate the profoundness of the entwinement of search engines and information in contemporary society.

Searching for content versus search as practice: Re-inflating information

Up to this point, much of our argument has been advanced by focusing on the use of search engines as deeply embedded into the social practices of everyday life in an obvious, self-evident, ordinary way, that is as an increasingly stable element in all kinds of routine practices, cooking dinner for friends, travelling as a tourist, school work, even being a patient, and so on. We think of searching for information as, on the one hand, having the purpose of digging up epistemic content that is needed to know something, as in the example that introduces the chapter, and, on the other hand, as something we just do, without ever actively reflecting on it. We want to develop an account of online information searching as an activity that is responsive to its situatedness within specific social practices *and* to its role in finding epistemic content that actually fulfils a type of need. Bernd Frohmann (2004) suggests framing this distinction in terms of two discourses in which the role of scholarly literature could be understood – a practice discourse and an epistemic

discourse. Frohmann argued for the significance of focusing on the former, while downplaying the relevance of the latter. The practice discourse is in his account predominantly interested in what is done with and through documents at various levels in society and how these practices together establish an hierarchical order of knowledge through citations, ranking factors, databases, knowledge organisations in libraries, and so forth. The epistemic discourse, on the other hand, focuses on how information conveys epistemic content. In Frohmann's understanding, the epistemic discourse of information is of little interest since, in his view, it is based on a flawed understanding of how science actually works. His aim in criticising the epistemic discourse of information, as already stated in the title of his book – *Deflating Information* – is to deflate the significance assigned to the notion of information in favour of doings with documents. It is tempting to approach search engines in a similar way – clearly the way their use is embedded into everyday life and society cannot be understood by narrowly focusing on information needs and the epistemic content each piece of information delivers, but requires accounting for the ways in which they are part of social practices. Moreover, when we search with a web search engine, we do not just get access to knowledge as content, we also participate in establishing an order of knowledge in which the data of searchers co-constructs what there is to find and know in the first place. In this way, together with other searchers' data, indexes, algorithms and so forth an order of importance is established (compare Rogers 2013).

The work of the cognitive information retrieval tradition on the evaluation of search engines and much research from the information behaviour tradition, including the concepts of information needs and relevance themselves, are founded within an epistemic discourse of searching, even if the content is mostly considered an abstract entity and not in its actual relation to social or political structures. Work on the politics of search engines and search engine bias does consider these entanglements, but it is in fact also founded within an epistemic discourse. In contrast, research on how searching is socially shaped is mostly based on a practice-based understanding of searching and here, as we have maintained throughout, information tends to be submerged within the various practices of whatever social situation is at stake. By *re-inflating information* our contribution is to make these three traditions visible and join them in order to better understand searching in everyday life. Re-inflating information, that is a balanced foregrounding of the constitutive entanglement of the notions of information as content and information as emergent in practice, is also necessary in order to be able to say something worthwhile about the categorisation of information and of people as well as of their treatment as commodities in how search engines work in contemporary society. Ola Pilerot (2014, p. 2013) has developed a similar notion of the double role of material objects in relation to information sharing, and advances an understanding of "objects as material artifacts with social (boundary) as well as mental (epistemic) features". If search engines structure the order of knowledge, and we claim that they do, then we need to understand how the information they provide is given meaning, in practices *and* in an epistemic sense. This becomes acutely obvious in

areas, which have a history of violence and oppression, as for instance Deirdre Mulligan and Daniel Griffin (2018) discuss in relation to the search query "did the holocaust happen" or as Safiya Noble's (2018) account of how racism is built into the very fabric of search engines, demonstrates at length. At the same time, the documentary shape of information in this context is prescribed by search engines and their largely invisible algorithms. Conformity with these only partially known rules is rewarded with exposure, disregard is punished with invisibility. From an advertisement perspective this makes sense. Yet, if applied to all types of information, also for instance to historic testimony or scientific facts, this leads to tensions that can be framed in terms of frictions of relevance, as the example of the results returned in response to the query "did the holocaust happen" makes so apparent (Mulligan & Griffin 2018). This has far-reaching consequences for what is an established fact and what is an issue of contention and on-going debate. As for instance the query "did the holocaust happen" – which now returns entirely different results than the ones originally subject to media coverage and public objection (Mulligan & Griffin 2018) – also shows, these arenas are moving targets. Furthermore, the effects of phenomena like Google bombing, that is making webpages show up high in response to certain, seemingly unrelated queries or the deliberate manipulation of autocomplete suggestions, can only be understood if we understand the meaning and wider implications of the terms concerned, in the cultural contexts in which they originate, but also in the cultural contexts into which they enter through searches carried out by users in their everyday lives.

In most instances online searching is actually a way of finding information representing epistemic content and often this content does have consequences. The story at the outset of the chapter clearly demonstrates this, but all kinds of political content, historic content, content reflecting cultural values or, of course, health and even product information are shaped by their relation to search engines and by how it is searched for. In this way, a narrow focus on practices falls short of elucidating what is at stake, as does a pure interest in the algorithmic shaping of content. Searching for information is often, but of course not always, carried out in order to relate to an information need, however submerged, at the same time as the act of searching co-constructs what there is to know in the first place. This can lead to conflicts between the individual, situational relevance, and societal interest. Earlier in this chapter we refer the potential friction between individual needs and experienced relevance on the one hand and societal interests and relevance for society as a whole on the other. Sometimes, this friction is made visible and reported by the media, such as when links are observed high up at the result page (despite the fact that we do not always actively seek for them) with blatantly false information or extremely offensive content. One reason could be the algorithm-conscious behaviour of certain groups that engage in Google bombing, exploitation of "data voids", or similar practices or the fact that extreme things in general fascinate us, even if we do not support or believe in them. We refer to these situations as forms of frictions of relevance that lead to infrastructural breakdown. However, not only are these frictions often individual, and can occur on

very different or even opposing grounds, they are also rare in everyday life. This became acutely obvious for instance when US president Donald Trump in 2018 accused Google of anti-conservative bias, in curious opposition to almost all previous criticism of the search engine as biased in more or less precisely the opposite way and characteristically also in defiance of most research on search engines and their various biases. At the same time, the friction of relevance Donald Trump's statement gave expression to constituted a form of infrastructural breakdown that found its way into the media and which for a fleeting moment exposed this otherwise invisible infrastructure to a great many users worldwide. We will return to this question in the following chapters, but for now we want to call attention to one of the difficulties involved in search engine ranking that earlier literature has also pointed out (e.g. Van Dijck 2013; Gillespie 2014; Halavais 2018[2009]; Hillis, Petit, & Jarrett 2013; Jiang 2014; Mager 2012; Noble 2018).

As we developed in Chapter Three, the way in which Google ranks the results of a search is a highly complex process involving Google's index, algorithms, and user data. Ultimately, practices such as Google bombing, conscious exploitation of "data voids" (Golebiewski & boyd 2018), and similar attempts – also officially sanctioned ones, as is not least common practice in China – to get search engines to order and present information in a certain way, can be thought of as attempts to actively shape discourse; discourse here understood in a loosely Foucauldian sense as what is possible to say and do. Search engines are implicated in laying down the rules for how truth is given meaning in the specific societal formations in which they are embedded. These discourse-shaping practices are often attempts to reduce the dissonance the above-discussed frictions of relevance provoke. They do this by, if not immediately normalising, then at least by digging out certain, often previously supressed connections and putting them on exactly the same stage as other connections that are socially and societally accepted. Metaphorically formulated, it is a gaming of the algorithm, but the algorithm seems to enjoy the game, at least it always appears to respond. The role of encyclopaedias can be understood in a similar way – as both providing us with epistemic information and constructing an order of knowledge and what to count as knowledge in the first place (Haider & Sundin 2014). Today, the most important and most widely available encyclopaedia is undoubtedly Wikipedia. Not just because it almost always ends up as the first link on a results page of Google search, but also – and in the future maybe even more important – because of how the content of Wikipedia is used when providing the searcher not just with a link, but with the direct "fact" through the Google Knowledge Graph function. Wikipedia and its free labour (Lund 2017) play an important role in how Google is changing from a pure reference database to a fact provider. Given that most searchers never even scroll down to the bottom of the first results page, it is not a far stretch to assume that for example shorter lookup searches in many cases stop at the Knowledge Graph, and thus – at least for the time being – at content produced by Wikipedia contributors. This implicates Google even deeper in the fascinating but also risky work of factualising information, glossing over the transparency that is characteristic of Wikipedia's model of peer-production.

Chapter conclusion: The ubiquity of search

Online search is a technically complex phenomenon where, under increasingly simpler surfaces, sometimes just a voice, hundreds of years of knowledge management, bibliographic tools, and information access and retrieval technologies are hidden, shaped to fit economic paradigms of capitalism, the rules of marketing, a complex media landscape, and far-reaching political issues ultimately concerning societal control over knowledge. While search is, of course, specific in each case, it is also the same. Clearly, googling in school is different from googling for a recipe or finding the backstory to a political scandal, and very different from searching for a funeral home or how to meet others grieving over the death of a relative. All these are embedded into networks of other information-related activities, analogue and digital, that are very different from each other. This difference needs to be taken seriously in how we conceptualise and talk about the role of information and the technologies used for finding it. Yet, they are also the same; they all use Google or another similarly dominant, corporate technology; in all cases not using it would be disturbing something that is expected and the way in which information is chosen and provided is increasingly left to proprietary algorithms under corporate control. We need to also acknowledge this sameness in order to be able to articulate the meaning and significance of the increasingly platformised infrastructure (Plantin et al. 2018) which structures access to information and shapes knowledge, on a societal level and also to critique it. In this sense, our account of search in everyday life wants to highlight the generalness of the specific, while showing how the specifics of certain situations bleed into other situations precisely through the use of the *same* technology. It also wants to balance different notions of information that come to the fore, that is information as something that happens and information as something that is – on one hand, information as emergent in practice and, on the other, information as epistemic content. The search engine needs both and in return both are shaped by the search engine. In order to understand and critique the methods for indexing, classifying, and evaluating content that make possible search as this specific mediated enactment of cultural values, biases, possibilities and constraints and its permeating everyday life, both understandings are, as we have tried to show, necessarily entangled in each other. That is, the constitutive entanglement of the material and the social needs to be situated – also – at the level of different notions of information.

References

Andersson, C. (2017a). "The front and backstage: Pupils' information activities in secondary school. "*Information Research*, 22(1), CoLIS paper 1604. Retrieved from http://InformationR.net/ir/22-1/colis/colis1604.html.

Andersson, C. (2017b). "'Google is not fun': An investigation of how Swedish teenagers frame online searching." *Journal of Documentation*, 73(6), 1244–1260.

Carlsson, H. & Sundin, O. (2017). "Searching for delegated knowledge in elementary schools." In Proceedings of ISIC, *Information Behaviour Conference, Zadar, Croatia, 20–23*

September 2016: Part 2. Information Research, 22(1), paper isic1618. Retrieved from http://InformationR.net/ir/22-1/isic/isic1618.html.

Carneiro, H. A. & Mylonakis, E. (2009). "Google trends: A web-based tool for real-time surveillance of disease outbreaks." *Clinical infectious diseases*, 49(10), 1557–1564.

Case, D. O., Andrews, J. E., Johnson, J. D., & Allard, S. L. (2005). "Avoiding versus seeking: The relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts." *Journal of the Medical Library Association*, 93(3), 353–362.

China Daily (2016). "Baidu's objectivity compromised by profit model." Retrieved from http://www.chinadaily.com.cn/business/2016-05/10/content_25184744.htm.

Collinson, P. (2018). "I've tried to dump Google and Facebook. But it's been painful." *The Guardian*. Retrieved from https://www.theguardian.com/global/2018/may/19/google-facebook-security-data-duckduckgo.

Cox, A. M. (2012). "An exploration of the practice approach and its place in information science." *Journal of Information Science*, 38(2), 176–188.

Frohmann, B. (2004). *Deflating information: From science studies to documentation*. Toronto: University of Toronto Press.

Gibbs, J. L., Ellison, N. B., & Lai, C. H. (2011). "First comes love, then comes google: An investigation of uncertainty reduction strategies and self-disclosure in online dating." *Communication Research*, 38(1), 70–100.

Gillespie, T. L. (2014). "The relevance of algorithms." In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds), *Media technologies: Essays on communication, materiality, and society* (pp. 167–193). Cambridge, MA.: The MIT Press.

Golebiewski, M. & boyd, d. (2018). "Data voids: Where missing data can easily be exploited. "*Data & Society*. Retrieved from https://datasociety.net/output/data-voids-where-missing-data-can-easily-be-exploited/.

Haider, J. (2017). "Controlling the urge to search: Studying the informational texture of practices by exploring the missing element." *Information Research*, 22(1), CoLIS paper 1613. Retrieved from http://InformationR.net/ir/22-1/colis/colis1613.html.

Haider, J. & Sundin, O. (2014). "The materiality of encyclopedic information: Remediating a loved one–Mourning Britannica." *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–10.

Halavais, A. (2018[2009]). *Search engine society* (2nd Ed.). Cambridge, UK: Polity Press.

Hillis, K., Petit, M., & Jarrett, K. (2013). *Google and the culture of search*. New York: Routledge.

Huvila, I. (2016). "Affective capitalism of knowing and the society of search engine." *Aslib Journal of Information Management*, 68(5), 566–588.

Jiang, M. (2014). "The business and politics of search engines: A comparative study of Baidu and Google's search results of Internet events in China." *New Media & Society*, 16(2), 212–233.

Johnson, J. D. (2009). "An impressionistic mapping of information behavior with special attention to contexts, rationality, and ignorance." *Information Processing & Management*, 45 (5), 593–604.

Kjellberg, S. & Haider, J. (2018). "Researchers' online visibility: Tensions of visibility, trust and reputation." *Online Information Review*. Retrieved from https://doi.org/10.1108/OIR-07-2017-0211.

Lewandowski, D. (2018). *Suchmaschinen verstehen* (2. Auflage). Berlin, Heidelberg: Springer.

Lloyd, A. (2014). "Following the red thread of information in information literacy research: Recovering local knowledge through interview to the double." *Library & Information Science Research*, 36(2), 99–105.

Lloyd, A. (2010). "Corporeality and practice theory: Exploring emerging research agendas for information literacy." *Information Research*, 15(3) CoLIS794. Retrieved from http://InformationR.net/ir/15-3/colis7/colis704.html.

Lund, A. (2017). *Wikipedia, work and capitalism: A realm of freedom?* Cham, Switzerland: Palgrave Macmillan.

Mager, A. (2012). "Algorithmic ideology: How capitalist society shapes search engines." *Information, Communication & Society*, 15(5), 769–787.

McKenzie, P. J. (2003). "A model of information practices in accounts of everyday-life information seeking." *Journal of Documentation*, 59(1), 19–40.

Marchionini, G. (2006). "Exploratory search: From finding to understanding." *Communications of ACM*, 28(9), 41–46.

Mavragani, A. & Tsagarakis, K. P. (2016). "YES or NO: Predicting the 2015 GReferendum results using Google Trends." *Technological Forecasting and Social Change*, 109, 1–5.

Mulligan, D. K. & Griffin, D. S. (2018). "Rescripting search to respect the right to truth." *Georgetown Law Technology Review*, 2(2), 557–584.

Myrick, J. G., Willoughby, J. F., & Verghese, R. S. (2016). "How and why young adults do and do not search for health information: Cognitive and affective factors." *Health Education Journal*, 75(2), 208–219.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

Ofcom (2018). *Adults' media use and attitudes report*. London: Ofcom.

Orlikowski, W. J. (2007). "Sociomaterial practices: Exploring technology at work." *Organization Studies*, 28(9), 1435–1448.

Pilerot, O. (2014). "Making design researchers' information sharing visible through material objects." *Journal of the Association for Information Science and Technology*, 65(10), 2006–2016.

Pilerot, O. (2013). "A practice theoretical exploration of information sharing and trust in a dispersed community of design scholars." *Information Research*, 18(4) paper 595. Retrieved from http://InformationR.net/ir/18-4/paper595.html.

Plantin, J. C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). "Infrastructure studies meet platform studies in the age of Google and Facebook." *New Media & Society*, 20(1), 293–310.

Purcell, K., Brenner, J., & Rainie, L. (2012). *Search engine use 2012*. Washington, D.C: Pew Research Center's Internet & American Life Project.

Rieder, B. & Sire, G. (2014). "Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web." *New Media & Society*, 16(2), 195–211.

Rivano Eckerdal, J. (2012). "Information sources at play: The apparatus of knowledge production in contraceptive counselling." *Journal of Documentation*, 68(3), 278–298.

Rogers, R. (2013). *Digital methods*. Cambridge, MA.: The MIT Press.

Schroeder, R. (2018). *Social theory after the internet: Media, technology, and globalization*. London: UCL Press.

Schroder, R. (2015). "Does Google shape what we know?" *Prometheus: Critical Studies in Innovation*, 32(2), 145–160.

Savolainen, R. (2006). "Time as a context of information seeking." *Library and Information Science Research*, 28(1), 110–127.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices." *Journal of Documentation*, 72(6), 990–1007.

Sundin, O., Haider, J., Andersson, C., Carlsson, H., & Kjellberg, S. (2017). "The search-ification of everyday life and the mundane-ification of search." *Journal of Documentation*, 73(2), 224–243.

Talja, S. & Nyce, J. M. (2015). "The problem with problematic situations: Differences between practices, tasks, and situations as units of analysis." *Library & Information Science Research*, 37(1), 61–67.

Taylor, R. S. (1968). "Question-negotiation and information seeking in libraries." *College and Research Libraries*, 29, 179–189.

Trias i Mansilla, A. & de la Rosa i Esteva, J.L. (2013). "Survey of social search from the perspectives of the village paradigm and online social networks." *Journal of Information Science*, 39(5), 688–707.

Tripodi, F. (2018). "Searching for alternative facts: Analyzing scriptural inference in conservative news practices." *Data & Society*. Retrieved from https://datasociety.net/output/searching-for-alternative-facts/.

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. New York: Oxford University Press.

Veinot, T. C. (2007). "The eyes of the power company: Workplace information practices of a vault inspector." *The Library Quarterly*, 77(2), 157–179.

Waller, V. (2011). "Not just information: Who searches for what on the search engine Google?" *Journal of the American Society for Information Science and Technology*, 62(4), 761–775.

Zavadski, A. & Toepfl, F. (2019). "Querying the Internet as a mnemonic practice: How search engines mediate four types of past events in Russia." *Media, Culture and Society*, 41(1), 21–37.

5

SEARCH AND MEDIA AND INFORMATION LITERACY

"Facts are created by measurement and these can be reviewed and repeated. Opinions only arise – like weeds" (Rosling 2017, translated to English by the authors). This sentence was written by Ola Rosling, director of the Gapminder Foundation and son of the famous Swedish health researcher, statistician, and public speaker Professor Hans Rosling who passed away in 2017. To a wider audience, Hans Rosling is best known for his popular and beautifully produced video presentations. Using the so-called Trendalyzer software, he visualised statistical trends and showed sometimes astounding connections in data available from open data sources. Bill Gates, co-founder of Microsoft, gave access to Rosling's posthumously published book *Factfulness* (2018) to all college graduates of 2018 in the USA. A recurring theme in the book and presentations is to demonstrate a *data-based* view of the world that often challenges widely held beliefs or prejudices which do not tally with the actual data we have available. The above quote likening opinions to weeds was part of an opinion piece – of all genres – that appeared in the largest Swedish daily newspaper in February 2017. It was one of many media contributions following the death of Hans Rosling. It also came at a time when buzzwords such as fake news, alternative facts, or filter bubbles were all over the media and the debate on the relationship between media, politics, and truth was carried on with a sense of rarely seen urgency. Hand in hand with the media interest in these issues went an increasing number of calls for the public to develop new skills to better assess the trustworthiness and credibility of online information, for media and information literacy, in other words.

Likening opinions to weeds, something to rid society of, as is done in the introductory quote, is a bold metaphor. It goes against a common understanding of what opinions are and how they relate to facts. Rather, in relation to data and statistical trends, opinions can be said to be interpretations of why the world looks as it appears through the lens of data tools or of how we should act in order to

DOI: 10.4324/9780429448546-5

change or support a certain trend, not seldom depending on the ideological ground that informs us. Also, opinions lie behind why certain types of data should be collected in certain ways in the first place and then how we should interpret them. This is how – for instance – political debate is sustained. Yet, in view of the way in which data are treated as factual representations of the world in the context of the Gapminder Foundation's work and in other similar contexts, it makes sense to position facts in opposition to opinions. This contrast is also constantly reproduced in certain areas of the public debate and it is within this public discussion that the above quote has to be positioned. Increasingly, if facts are criticised, which is seldom done by brushing them aside as being mere opinions, while the media's various opinion sections have come under scrutiny for not being objective or neutral enough, for not being factual in other words. So, what does this muddled up antagonism tell us about the roles different types of information are assigned in contemporary society, at school and in everyday life? And what does it have to do with media and information literacy in relation to the use of web search engines?

The central theme of this chapter concerns the difficulties involved when attempting to turn search and searching into objects of learning and teaching, and how this plays out at a time of intense public debate about the nature of facts, disinformation and the role of platformised information infrastructures in this. We turn our attention to implications these difficulties have for developing certain aspects of media and information literacy instruction as well as for researching media and information literacy. While search is a mostly invisible, taken for granted use of software tools submerged in different devices and everyday life, and as such often seen to be neutral, search engines are – as we have argued and shown throughout – value-laden at all levels. The problem we are facing boils down to the question, how can these critical aspects be made visible, without making searching with general-purpose search engines more or less impossible by continuously questioning the trustworthiness of all search results? This dilemma is related to a fundamentally epistemological question that touches upon the balance between how we know and what we know and how we articulate criticism of various epistemological assumptions without entirely dismissing their legitimacy. It is a question of how society can respond to what Bruno Latour describes as a form of "instant revisionism" (2004, p. 228), that is when established knowledge of, for example, environmental research or research on climate change, is questioned by reducing everything to a mere question of perspective often in a vulgar interpretation of postmodern relativism. It is an issue that all media and information literacy research and practice has to engage with. There is no definitive answer or solution, but the question of how can media and information literacy incorporate trust in knowledge institutions without giving up its critical stance needs to be posed, and it needs to be related to this form of "instant revisionism" and other attempts at vulgarising constructivist epistemology for the purpose of turning it against itself, in the course disarming all forms of constructive criticism.

Media and information literacy

Media and information literacy, often abbreviated to the policy-friendly MIL, combines two fields – media literacy and information literacy. These two fields are not just fields of research, but also fields of policy-making and professional practice. Research in media literacy has its home in media studies, while information literacy research has a background in library and information science. Their coming together, mainly during the early 2000s, is a complicated story where different research traditions and policy demands blend and contradict each other. "Information literacy research has attended more to questions of access, while media literacy research has paid more attention to questions of understanding" write Sonja Livingston, Elizabeth van Couvering and Nancy Thumin (2008, p. 108) in an influential chapter on the converging traditions of research on media and information literacies. The notion that media literacy focuses on understanding and information literacy on technology and access is accurate, yet only to some extent. That story has mostly been written from the perspective of media literacy research and media studies. In it, the critical and theoretically more nuanced approaches existing in library and information science have often been neglected.

There exists in fact a tradition in both research and professional practice that criticises the more instrumental approaches to information literacy and their focus on access and reduction of literacy to a largely context-independent set of skills. Lifting such more nuanced approaches into the compound media and information literacy is more productive than merely focusing on those prioritising the technical side of things or issues such as access. Specifically regarding search and search technology, this becomes very obvious, since neither access to nor skills in using search engines are really what pose problems. Rather, what is at stake here are processes of meaning making, yet – considering particularly the information side in MIL – less related to the media sources linked to on the results pages, and more of the search results themselves and of how they came about. This can be understood, we suggest, as a type of *infrastructural meaning-making*. Access is rarely a straightforward issue and understanding access is part of this meaning making. In other words, embedding search into the notion of media and information literacy makes clear that it needs to incorporate infrastructural meaning-making in order to account for the various platforms, such as search engines, through which media and information literacy is enacted.

Information literacy research

Here, it makes sense to revisit notions of information literacy, which, compared to media literacy, has been given less attention outside its own field of application, and we begin by calling the attention to *critical information literacy* (e.g. Elmborg 2006; Hicks 2013; Kapitzke 2003; Tewell 2015, 2016). What characterises this, what we can almost refer to as a movement, is a common ground in critiquing so-called list approaches to information literacy and the definition of information literacy as it is provided by the American Association of College and Research

Libraries (ACRL) where – according to the criticism articulated – information is seen as neutral and literacy as sets of mostly technical skills. Often informed by Paulo Freire's work on critical pedagogy, the focus tends to be on learning in higher education and librarians are positioned as active agents offering students tools for a form of ideological resistance within a largely capitalist information and education market. There are also other examples where critical information literacy constitutes part of a more generally framed critical literacy, in particular regarding how librarians and other information professionals enable multifaceted, responsive approaches to embed critical awareness in different types of engagements with information, various cultural expressions and the media (McNicol 2016).

The critical information literacy movement comes often back to, examines and questions the traditional role of the librarian as a neutral mediator and emphasises how information and information access are socially constructed (Kapitzke 2003). James Elmborg (2006), for instance, stresses the need for librarians' information literacy instructions to be based on an understanding of what information is regarded to be in academia – the academic discourse – as well as in a critical understanding of what a library is and the role of librarians. In his extensive literature review of the topic Eamon Tewell (2015) underlines the socio-political role of librarians by stating that "[c]ritical information literacy considers in what ways librarians may encourage students to engage with and act upon the power structures underpinning information's production and dissemination" (ibid., p. 25). Tewell (2016) draws on resistant spectatorship theory when he suggests how librarians can support students in developing a critical understanding of Google Search. In this attempt, he questions the positioning as neutral of both librarians and search engines. The librarian becomes a political actor who should help reveal the biased and commercial character of Google Search, including its algorithms, thus giving rise to a richer notion of information literacy that accounts for the corporate structure of search technology.

Concurrently, a similar understanding of information literacy has been advanced by others, outside the USA; equally critical of simplified list approaches, where information literacy is a freely floating skill that can be acquired and then applied independent of its context. Veronica Johansson (2012, pp. 48–55) highlights the influence of so-called New Literacy Studies (NLS) for the development of a nuanced, richer picture of literacies in the plural as enacted in relation to specific situations and practices. In parallel, practice theory in different flavours made its entrance and profoundly transformed the study of information literacy, on the one hand situating it in specific practices, on the other hand moving it into contexts beyond traditional educational settings and removed from these institutions. Our concern in this chapter is however with what could be considered traditional arenas for education and the issue of how to turn online search and search engines into objects of teaching and learning in these settings and in relation to these newer, nuanced notions of information literacy as critical and multiple. In Nordic research on information literacy, with an empirical focus on compulsory school and upper secondary school and often informed by a sociocultural and Vygotskian tradition of

learning, a similar agenda has been developed of improving students' analytical abilities in relation to information (e.g. Limberg et al. 2008; Sundin & Francke 2009). Kimmo Tuominen, Reijo Savolainen, and Sanna Talja (2005) advance an understanding of information literacy as a sociotechnical practice. In agreement with the critical information literacy movement, they emphasise how technology cannot be regarded as neutral and that information as well as activities should be understood as relative to communities of practice. Such a sociotechnical position can also be found in Annemaree Lloyd's work, but, rather than emphasising the affordances of technologies, she argues for considering bodies and corporal experiences (e.g. 2014). As we stress in Chapter Two, there seems to be almost an inner contradiction between a practice approach and an interest in information searching. Thus, in this research tradition, information literacies are – and we agree with this position – seen as being intimately connected to the specific situation from which they emerge. Yet they are rarely understood as something that – also – means employing tools, which are distinct, recognisable information retrieval systems and which do in fact work similarly across different practices and situations, thus potentially entangling those situations into each other.

A research interest in the evaluation of online information can be found in information literacy and neighbouring fields of research in relation to both school and everyday life. A number of related concepts are used in this line of research, such as credibility, believability, reliability, and trustworthiness (e.g. Savolainen 2007). In this book, the term critical evaluation of information is used to describe the activity of evaluating if certain information (and information sources) is to be trusted *enough* in order to be acted upon. That could in contemporary online environments be, for example, a website, a book, a Facebook status, a blog post, a tweet or – which we focus on – the workings of an online search engine, such as Google, in order to challenge the given presentation of results and the order of ranking if necessary. In information science, the concept of cognitive authority (Wilson 1983), that is how we come to trust certain people, institutions or things, has often been used to analyse critical evaluation of information (e.g. Francke & Sundin 2009; Rieh 2002; Savolainen 2007). People tend to find information credible when it can be related to what is considered a cognitive authority, for example an author, a publisher, an institution or – as we will see later – a platform brand name or even an algorithm providing order to a search.

Mikael Alexandersson and Louise Limberg (2013) show how the development of, in their case, education in Swedish schools since the 1990s has encouraged various forms of independent learning and thus increasingly required independent information searching. A similar didactic development is likely to be found also in other countries. Students at almost all levels are often asked, ideally with guidance from teachers, to find their own school material, their own information, in order to complete their school tasks; the monopoly of the textbook has been challenged. As a result, Alexandersson and Limberg (ibid.) argue, the focus of those studying these developments and pedagogical methods has been placed on pupils' *evaluation* of information. This has given rise to a plethora of studies on evaluation of

information (Limberg et al. 2008). It is not just a growing interest in evaluation of information as a research topic, but abilities to evaluate information have also been given increasing attention by librarians and teachers, as well as been emphasised in the public debate (as we will return to below). While this has led to important contributions and debates, we argue it has also drowned out interest in information searching in educational settings.

There are different responses from society to the growing appearance and awareness of what are called fake news, alternative facts, or more accurately misinformation and even propaganda. Leaving aside the responses by authoritarian governments, in democratic countries there is a growing interest in fact-checking services and tools and these have begun to proliferate. Another response can be found in governments' legal actions and their pressure on platforms to exercise self-cleaning, for example in Facebook's handling of troll accounts or Google's handling of links to racist material. Sometimes changes occur in reaction to unfavourable media reports or to law suits brought in by citizens. Our focus here is on a further response, namely the demand and growing responsibility put on individuals to be able to evaluate online information on their own, which is captured in calls for increased media and information literacy and for improved skills to enable critical evaluation of information. In the public debate, as exemplified by Ola Rosling's quote mentioned earlier or by the proliferation of fact checking services, evaluating information, and information sources often simply means verifying whether content is correct or not. Sometimes this is enough. Yet in many cases, this is not what is at stake. Consider the following quote from a media interview with Google's search liaison, Danny Sullivan, appointed to the role in 2018:

> "We're not a truth engine. One of the big issues that we're pondering is how to explain that our role is to get you authoritative, good information, but that ultimately people have to process that information themselves," he said. "We can give you information, but we can't tell you the truth of a thing".
>
> *D'Onfro 2018*

The substance of this striking claim is not particularly controversial. What is interesting is that it is made in the first place and that the question of truth is touched upon at all, albeit only to be transferred to the realm of the user. Almost ironically, after two decades of advancing itself as a neutral broker of information (Hillis et al. 2013) that would enable the type of fact checking that is increasingly demanded, Google itself appears to become wary of its own creature, calling – implicitly – for improved media and information literacy.

Media and information literacy in policy making

The media and information literacy concept – MIL – has been promoted by actors such as UNESCO, IFLA, and numerous other similar international and national organisations and interest groups. Media and information literacy has often been

advanced by various actors as a means of reinforcing citizens' abilities to participate in society. UNESCO, for instance, phrases this idea on its website in the following way:

> Media and Information Literacy recognizes the primary role of information and media in our everyday lives. It lies at the core of freedom of expression and information – since it empowers citizens to understand the functions of media and other information providers, to critically evaluate their content, and to make informed decisions as users and producer [sic!] of information and media content.
>
> *UNESCO 2017*

For UNESCO, media and information literacy is an active subject. It acts on its own; it "recognizes" and it "empowers" individuals to make decisions. The ability which is advanced here to contribute with (publish), search for, critically evaluate, and use information in different practices is often framed as a basic condition for a democracy where citizens' active participation and dialogue is a prerequisite. Such an understanding of democracy demands citizens who have the intellectual and technical prerequisites to take part in a society infused with digital technology; it advances the ideal of the informed, responsible and reasonable citizen, who competently makes rational decisions (see also Pilerot & Lindberg 2011). When in May 2018 the Swedish National Commission for UNESCO gathered together ministers, politicians (no fewer than four ministers from the cabinet), scholars, and authorities to discuss the collaborating efforts in relation to media and information literacy in society, it was a typical example of a blend of policy, professional practice, and research. The invitation letter had the heading (translated into English) "Media and Information Literacy – a Question of Democracy" and the context was described in the following words: "The demands on the individual to critically evaluate information increases in the digital media landscape. The question of media and information literacy (MIL) is therefore today urgent" (UNESCO 2018). Promoting media and information literacy has been regarded as one of the ways society takes on the consequences of everything from misinformation to more generally preparing people for the almost non-existent barriers to publishing and the massive amount of information on almost any topic.

A growing interest in facts, trust, and the role of education is a global phenomenon, just like the problem that media and information literacy is supposed to solve. The director of the OECD's Directorate for Education, responsible for developing the influential PISA test, stated, according to the BBC in a talk in March 2017, "[i]n the past, when you needed information, you went to an encyclopaedia… and you could trust that the information would be true" and he continued: "Distinguishing what is true from what is not true is a critical judgement" (Coughlan 2017). Since 2018, PISA, for those countries that wish to participate, measures pupils' so-called "global competence" and one aspect of this competence is described in the following way:

> Globally competent students are able to **reason with information** from different sources, i.e. textbooks, peers, influential adults, traditional and digital media. They can autonomously identify their information needs, and select

> sources purposefully on the basis of their relevance and reliability. They use a logical, systematic and sequential approach to examine information in a text or any other form of media, examining connections and discrepancies. They can evaluate the worth, validity and reliability of any material on the basis of its internal consistency, and its consistency with evidence and with one's own knowledge and experience. Competent students question and reflect on the source author's motives, purposes and points of view, the techniques used to attract attention, the use of image, sound and language to convey meaning, and the range of different interpretations which are likely for different individuals.
>
> *OECD 2018*

To select and critically evaluate information is here framed as the responsibility of everyone, everywhere, and media and information literacy is consequently promoted as school content. The quote from the OECD is an example of a growing anxiety among official authorities and organisations about the lack of abilities among citizens to handle the changing digital environment. The quote also exemplifies something else – the invisibility of search. The ability to search is here taken for granted and we will get back to that later on, but first we would like to dig into the concept of the critical evaluation of information.

How do we evaluate information?

Information literacy research has many times investigated how people evaluate information, often with reference to cognitive authority. As indicated earlier in the chapter, most often research shows people's lack of abilities. Also our research (e.g. Sundin et al. 2017; Sundin & Carlsson 2016) shows that we often do not make the rational critical evaluation of information that policies, research and professional practice frequently hope people would make. At the same time, it would simply not be possible if we always had to question the trustworthiness and credibility of information we come across in our daily business. While credibility is here seen as what people consider as being likely or even true in relation to information, trust is seen as referring to whether people believe in the person, institution, publisher, and so forth providing the content. A society in which trust is supplanted by the need for posterior evaluation of all information is a world where public debate is rendered impossible. We take as a starting point that knowledge is intertwined with trust (Hardwig 1991). In order to accept something as knowledge in the first place we have to have trust other people (interpersonal trust), but in addition also institutions, technology, systems, processes and so on. In order to come to terms with how knowledge is communicated in contemporary online environments through the means of databases, search engines and other resources for information, Judith Simon (2010, p. 346) describes the trust/knowledge-link on the web as sociotechnical epistemic system and states, "we do not only place trust in human epistemic agents, but also in non-human agents, in epistemic processes as well as in epistemic content itself". Most of the time when using web search engines, we in fact put trust in the capacity of the search engine to deliver relevant information for us, rather than only

trusting the individuals who have produced the content provided through the links from the search engine results page. Simon also refers to the notion of algorithmic trust, originally suggested by Clay Shirky (2009) as a way of understanding how we tend to trust all kinds of filtering devices on the web that provide us with sorted and ranked information according to the principles of the algorithms involved. Simon (2010, p. 354) concludes, "[t]o be responsible knowers we must not only be willing to assess whether we are warranted in trusting epistemic content and its providers. We must also make our methods for assessing trustworthiness subject to scrutiny". We take this notion of trust with us in what follows.

Social epistemologist Dario Taraborelli (2008) introduces the notion of epistemic deference to describe how people's trust in information when forming an understanding is based on new information. He distinguishes between evaluative judgement and predictive judgement when examining the evaluation of information. Evaluative judgement refers to the evaluation of certain content while predictive judgement concerns the evaluation of a source made in advance of accessing the content. Taraborelli argues that too much research effort has gone into investigating people's evaluative judgement while predictive judgement has often been ignored. That claim could also be related to how evaluation of information is investigated in information literacy research. To break this down, it is difficult to evaluate a content separated from its medium. Take, for example, pictures that can easily be manipulated in a way that it is not possible for most of us to see through, a situation that the emergence of deep fake videos is only going to worsen. Predictive judgements, on the other hand, "are those that users make when they evaluate a source on the basis of information describing a source, like in the case of link descriptions" (Taraborelli 2008, p. 197). Google's and other search engines' *results snippets* function is an example of information describing a source. It is interesting here to return to the distinction between analysing searching with an interest in practices versus doing so with an interest in epistemic content that we outlined in Chapter Four. Taraborelli's notion of predictive judgement can here be understood as informed by a practice approach, while evaluative judgement can be understood in relation to an approach foregrounding epistemic content.

From information literacy research, we learn how pupils tend to make a sharp distinction between facts and opinions and how facts and opinions often, in the eyes of pupils, are tied to genres (Francke & Sundin 2009). Related to credibility evaluation the question of "facts" is of importance and a fact is in these situations often regarded as a core substance upon which the understanding in certain knowledge domains is built. Yet what exactly are facts considered to be in school situations more precisely and how are they imagined? Cecilia Gärdén and her colleagues (2014) conclude from a synthesis of four different research projects that facts – at least in Swedish school settings, an this can most likely also be applied to other similar contexts – are attributed three distinctly different meanings: "facts according to genre and modality", "facts as concrete external entities" and "facts as true and neutral". As a term, "fact" has a central presence in the classroom (Limberg 1999), and there is some evidence that search engines re-construct otherwise

complex issues as searchable, thus simplifying them to fit the affordances of the search engine (Rieh et al. 2016). The introduction of search engines in the classroom can be argued to thus contribute to a growing emphasis on "facts" at the expense of "understanding". Soo Young Rieh and her colleagues (2016) even claim that "current search engines are optimized for only a certain kind of learning – acquiring factual knowledge" (p. 20). The authors argue for supplementing Marchionini's typology (2006) of search with a concept they call *comprehensive search*, and which they describe as consisting of "a variety of search activities that facilitate individuals' critical abilities and/or the development of new ideas directly" (p. 23). Compared to Marchionini's notion of exploratory search, comprehensive search, according to Rieh et al. (2016), is, with a starting point in cognitive constructivism, more oriented towards fostering creativity. They differentiate between two types of knowing, firstly "the ability to concentrate on one thing deeply for a long time" and secondly "the capacity for explorative and integrative thinking" (ibid., p. 25). *Look up search,* another of Marchionini's types of searches, is supposed to support the former while *comprehensive search* is seen to support the latter. The authors go as far as defending what some contemporary critics have observed as shallow learning with the argument that comprehensive, less focused, search is better suited to the digital environment for learning.

Social media and search engines are important gatekeepers for news media (Mitchell, Jurkowitz, & Olmstead 2014). At the same time, people's abilities to handle these tools are not always strong – when it comes to certain aspects of infrastructural meaning-making. For example, a report from the USA shows the difficulties middle school pupils experience when trying to distinguish commercial ads from other web content provided by social media (Wineburg et al. 2016). A similar result is found for German internet users when they assess search engine results (Lewandowski et al. 2018). Relatedly, a survey of adults in the United Kingdom shows that only 54% of those surveyed were knowledgeable about how Google Search is mainly financed and for YouTube the figure is 45% (Ofcom 2018). Together, these findings demonstrate some of the difficulties involved when assessing and understanding search results or social media feeds for that matter that are not directly related to "skills", as in being able to provide the engine with query terms. Sam Wineburg and colleagues (2016, p. 4) summarise their findings as follows: "Overall, young peoples' ability to reason about the information they found on the Internet can be summed up in one word: *Bleak*". The authors do not seem to indicate that adults' abilities are more developed, just that abilities for reasoning about the trustworthiness of online information do not necessarily follow experiences in using social media or other contemporary online information services in general. There seems to be a misconception based on a mix-up between technical skills and what we name infrastructural meaning-making. Infrastructural meaning-making also challenges the established distinction between search as skill and use as intellectual evaluation of results (sources), as processes following each other and brings them together at the same level.

Astrid Mager (2009), in a study of use and provision of health information online distinguishes between how, due to Google's re-arranging of information, users tend to be invested in a what she calls "issue-centred information epistemology" while site producers are grounded in an "actor-centred information epistemology". While the latter takes a starting point in the totality of one site, users construct a totality out of a number of sites. The user then becomes very dependent on Google and its ranking, without always knowing how it works. Communication researchers Miriam J. Metzger and her colleagues (2010) remind us that credibility evaluations are most often profoundly social and that people tend to evaluate information by social heuristics through, for example, considering its reputation, endorsement by others or comparing it with other websites. Typically, only when the motivation to evaluate a specific claim is high, a careful evaluation of online information is undertaken (ibid.). With low motivation, it is more likely that heuristics will be used (Metzger 2007; see also Höchstötter & Lewandowski 2009; Pan et al. 2007). As repeatedly noted throughout, when searching for information, particularly in everyday life, people tend thus to rely on the algorithms of search engines, and mostly of Google, rather than carrying out a critical evaluation of the retrieved information.

It seems that we have to broaden our focus; from a strict focus on critical evaluation of information or even information sources to include an understanding of the various paths we use to get the information we get, what we might consider in terms of a critical evaluation of search as a form of, as we call it, infrastructural meaning-making.

Understanding search versus understanding with search

Most people regard the search for information in general purpose search engines as something easy to do. For instance, a Pew report from 2012 reports that 56% of search engine users in the USA are very confident in their abilities to use search engines, while 37% are somewhat confident (Purcell, Brenner, & Rainie 2012, p. 13). For users under 50 years of age, the figure is 64% who are very confident in their abilities. In a report from the Swedish National Agency for Education (2016), it is shown that 87% of the pupils in upper secondary schools believe they are good or very good at searching for information. The percentage for those who believe they are good or very good at critically evaluating information they have found on the internet is lower, but still as high as 75% (ibid.). The above figures are supported by Melissa Gross and Don Latham's (2012) research on the relation between actually performed information literacy skills and self-view of these abilities. They show that the self-view college students have of their abilities does not tally with their actual skills. The point here is not to crunch numbers and compare small statistical differences, yet what these and similar reports demonstrate is a gap in peoples' experiences of their understanding of search and their critical evaluation of information and how teachers and researchers understand their abilities. How can we motivate a critical evaluation of searching as a type of infrastructural meaning-

making at a time when searching is seen as something simple, often as a lookup search? People tend to outsource critical evaluation of information to the algorithms of search engines (Sundin & Carlsson 2016) and this has implications for how instruction on the workings and use of general-purpose search engines is integrated into media and information literacy education in school settings. A number of the focus groups that we presented in Chapter Four consisted of teachers. In the discussions, they rarely identified searching as something their pupils have difficulties with or more generally as an object of learning, not unlike how searching is invisible and taken for granted in OECD's notion of "global competence" mentioned earlier. Instead, they most often talked about searching as a straightforwardly simple and practical skill. At the same time, the teachers witnessed how their pupils were dependent on the Google ranking: "The pupils buy the first thing they find". The absence of an ability for critical evaluation goes, according to the teachers we talked to, hand in hand with the pupils' unreflected trust in Google. The teachers also identify the rush they find their pupils in as a problem. Lack of time propels the outsourcing of critical evaluation to Google's relevance judgement even more. The issue is thus that the outsourcing of critical evaluation of information is not grounded in an informed understanding of the workings of search engines.

The layout of search engine results pages has significant bearing on where users are most likely to click (Höchstötter & Lewandowski 2009; Kammerer & Gerjets 2012, 2014), as does the location of paid results on these pages (Lewandowski et al. 2018) as well as the order of links (Pan et al. 2007). All these lead users to preferably choosing links shown high up on the results page. Eszter Hargittai and her colleagues (2010) summarised their research on young adults, more precisely US-American students, in a familiar air:

> To complete many of the assigned tasks, students often turned to a particular search engine as their first step. When using a search engine, many students clicked on the first search result. Over a quarter of respondents mentioned that they chose a Web site because the search engine had returned that site as the first result suggesting considerable trust in these services.
>
> *Hargittai et al. 2010, p. 479*

In other words, as we also showed in Chapter Four, people do not – at least not on a routine basis – thoroughly assess the information or the source in which the information is located. Not even in school contexts can the critical evaluation of information be taken for granted. Relatedly, Heidi Julien, and Susan Barker (2009), in their investigation of the conformity between the attention Canadian curricula pay to information literacy and high school students' actual abilities, argue that the growing awareness in the curricula was not followed by a development of students' actual skills. Instead, very high trust is put in the search engine's actual ability to provide the searcher with the "best" information. Again, it seems as if the trust in algorithms is a more important phenomenon than trust in the author, or even trust

in the epistemic content (see also Simon 2010). Drawing on the concepts developed by Taraborelli (2008) introduced above, it seems as if in web search predictive judgments have more weight than evaluative judgements (compare Huvila 2013).

In many countries Google has for a number of years been a self-evident companion in the classroom. This was preceded by a growth in student centred-education, which here refers to a way of teaching that creates space for students working independently or in groups with self-directed learning (e.g. Limberg et al. 2008). In these situations, the teacher becomes more akin to a supervisor than a content mediator. A condition for these didactics is a view of learning resources in the plural rather than relying on a single authoritative textbook. There is a debate on the pros and cons of student-centred education, and most teachers do, of course, combine different teaching methods. Yet, at least from a Western perspective, it is safe to say that independent learning is more common today than it was 30 years ago. A core activity in students' independent learning is to find relevant and reliable information. In a situation where the amount of information available on almost any topic is inconceivably large and for many purposes fast and easy to retrieve, the question then becomes: is there a need for knowing anything more than how to find the information you need or have we gone from learning just in case to searching just in time? The answer to the first part of this question is, evidently, yes, but it also raises awareness of how search engines to some extent have come to function as external memories, despite obvious problems with version control and with how they function as archives for their own past (Hellsten et al. 2006). The second part of the question is more difficult to answer in the present context. We cannot in a society as complex and interconnected as ours know everything in advance, but at the same time we need to have a general understanding – an all-round education – so as to be able to contextualise and interpret what is found on the web. In school, at the workplace and in everyday life more generally, we rely on the availability of finding information through searching online.

Media literacy scholar David Buckingham (2006, p. 263), with a nod to Umberto Eco, postulates that "education *about* media is an indispensable prerequisite for education *with* or *through* media". If we translate that proposition to search, that would mean that education about search (and search engines) is a precondition for education with or through search. Put differently, we need to make search visible not just as a means to an end, but as the object of learning in its own right (Sundin & Carlsson 2016; compare Limberg & Folkesson 2006. Search engines are media that organise other media. That said, to make online searching visible and into the object of learning does by no means of necessity lead to an argument to separate teaching about search from other subjects taught in school; to treat media and information literacy independently. On the contrary, search can be made visible even if media and information literacy is positioned as necessarily integrated into other subjects. Still, in order for the concept of media and information literacy to be meaningful, it must have some kind of content independent of the subject or practice of which it is a part of.

For teachers it is a challenge to make search in particular, but also many other aspects of information literacy, into the content of teaching and learning (Sundin & Carlsson 2016; compare Limberg & Folkesson 2006). Thus, and not surprisingly, search and searching are – by teachers we have talked to – most often referred to as neutral tools for finding information, and the focus has been on the information as such, once found, on its content or credibility, rather than on searching in all its complexity (ibid.). However, in order to come to terms with the way in which search engines function as information infrastructures constitutively entangled across practices, there is a need to develop a broader understanding of search, an understanding that not just considers search as a narrowly viewed technical skill.

While the dominant framing of information searching in school contexts seems to be focused on technical skills, thus flattening much of the complexity of searching for information, other understandings also circulate and give it meaning. Considering these helps us to articulate some of the difficulties at stake that we touch upon above. For instance, bringing in an example from the Swedish curriculum, it is clear that in these documents different notions of information searching co-exist (Sundin 2015). In one understanding, information searching is identified as the activity of finding information from one source, such as being able to read an article and extract the most important message. This can be described in terms of *intra-textual information searching*. A second understanding of searching concerns the activity of searching for information in a huge collection of documents, such as when googling. This can be termed *inter-textual information searching*. What we are most interested in here and what is in fact at stake in infrastructural meaning-making is how these two notions relate and constitute each other.

When, for example, information searching or similar activities are tested in international tests such as PISA, ICCS, or PIRLS, it is intra-textual information searching that is tested. Inter-textual information searching cannot easily be tested in the way standardised tests need to be done. The idea of letting students solve tasks and search for information independently does not fit standardised tests and especially not those that compare very different countries indeed. This is probably one explanation why search and the infrastructure for search is often absent from teaching. The distinction between intra-textual and inter-textual information searching could be related to what Sam Wineburg and Sarah McGrew (2017) label as reading "vertically" and reading "laterally". Reading vertically is the traditional way of approaching a text, where you start at the top and read through to the end. Reading laterally, on the other hand, refers to how students critically evaluate information by comparing different texts, for instance through opening up new tabs in the browser. Wineburg and McGrew (2017) maintain that professional fact checkers read more laterally, while teachers and university students read – with lesser success compared to the fact-checkers – more vertically. It seems as if a new kind of reading, more lateral and closely related to searching, is developing.

Changing demands: The limitations of media and information literacy

In the introduction to this chapter, we talked about an opinion piece articulating a wish for a clear-cut distinction between facts and opinions, between the good and the bad, measurable facts and worthless opinions. There is plenty of evidence that contemporary information infrastructure can provide each and every one of us with the resources to find information – often claimed to be facts – as proof of almost any statement or for a version of the truth. Andrejevic (2013) argues very convincingly for this in his book *Information Glut*. We – not least, of course, as academics in the social sciences and humanities – have effectively learnt how to puncture all kinds of knowledge claims and how to show how facts are made, how knowledge is produced, and discourse develops. Distinguishing a fact from what has come to be called an alternative fact in social media and search engines, or news from disinformation, can for sure be done in some cases, particularly when we talk about totally made up statements, but many factual and news claims are not of that totally invented type that will easily be revealed.

Bruno Latour (2004) reminds us in "Why has critique run out of steam? From matters of fact to matters of concern" of the need to reconsider our methods and the role and ways of being critical (see also Andrejevic 2013). More to the point, the challenge is, as he develops at length, that the same type of criticism of how a fact comes into being can also be employed to dismantle it. In his paper, Latour takes on the task of rephrasing his critical aim as a scholar – from focusing on matters of facts to focusing on matters of concern. Similarly, rather than focusing on facts, we should, he argues with reference to Martin Heidegger, pay attention to *Things* and how they come into being. He carves out an alternative position to what he refers to as the *fact* and the *fairy* – a *fair* position. Such a position offers ways "to detect *how many participants* are gathered in a *thing* to make it exist and to maintain its existence" (Latour 2004, p. 246).

It could be argued, admittedly glossing over a number of complexities, that search engines, their algorithms, as well as the production and circulation of data, contribute to disguising some of the various gatherings at work for issues to take shape as factual. What ends up at the top of the list following a Google search or what is returned by the voice assistant is what most people regard as accurate, as factual and for many types of searches, this can be a very rational course of action. Nevertheless, search engines also decontextualise information and, in this sense, they hide how issues are shaped through the workings of search engines or, more correctly, through how search engines are understood by large parts of the general public as neutral tools for retrieval of facts. This could be the scholarly understanding of how vaccines work, reasons for and consequences of climate change or – let us say – the presumed existence of chemtrails in the sky, to refer to a popular conspiracy theory. Rather than only focusing on evaluation of information in a binary way, as false or true, more emphasis needs to be put on understanding

how the information infrastructure of ours co-constructs trust and credibility by ranking and filtering news and knowledge claims.

Latour (2004, p. 246) also underlines, and this is important for our view of media and information literacy, "The critic is not the one who debunks, but the one who assembles". Rather than predominantly focusing on a critical evaluation of information, once retrieved, and seen as more or less decontextualised from its production and from its infrastructure, we need to investigate these infrastructural arrangements in order to understand why certain information is found for us and why other information is not, how this differs in different subjects and how this is carried out within different practices. How do people, institutions, data, indexes, algorithms, software, subjects and practices hang together? More theoretically speaking, we need to analytically understand the sociomaterial assemblage of search in order to carry out a critical evaluation of information understood as *infrastructural meaning-smaking*, and that means a productive rather than a merely destructive act.

One theme in the focus groups with teachers concerns the downside of the strong focus on *critique* inscribed in the very notion of critical evaluation of information. In a discussion on how the teachers try to promote a critical stance, one teacher reflects: "But at the same time, you leave them with a 'you-can't-trust-anyone-world' and that's not good either". Another quote illustrates the same problem. A teacher talks about her experience with pupils who "don't trust anything since they have learned to be good at being critical towards information". Thus, if the pupils have learned their lesson and question all knowledge claims, then essentially, the school's version of knowledge is no better than anyone else's. It is one thing to promote critical evaluation of information, but it is something else to provide pupils with tools to value and compare diverging knowledge claims. In a society dependent on second-hand knowledge trust is essential. If trust in traditional knowledge institutions such as schools, universities, and media vanishes and knowledge increasingly is framed in terms of a personal choice instead of a societal, a collectively negotiated value, this is a challenge that media and information literacy, as something tied to an individual, cannot tackle. Interestingly, the framing of knowledge as a personal choice that circulates in the public discourse appears to have its roots in a deliberate misrepresentation or at least a vulgar and distorting reading of constructivist positions. In this tradition, production of knowledge is analysed in relation to how power operates in society and as a deeply social process and not in any way "subjective" or "personal" as the populist criticism continues to maintain. Yet, critical evaluation of information and parts of media and information literacy are often taught in ways that foreground personal responsibility and individual choice and often for good reason. Choosing a link from a search engine results page is in many ways the epitome of this act. Blending these two strategies makes for a very difficult marriage; one requires – at the least – considerable historical knowledge and an awareness of societal developments, of power structures and institutions, of epistemology, and so on and the other one is often called upon to solve immediate problems of fact checking, of true or false, credible or not, confirming a certain narrative or not. In other words, making visible the social

construction of knowledge claims and evaluation of information get mixed up into each other. Yet the parts that are often brought together are not necessarily the ones that are most productive together. Thus, increasingly knowledge claims are being dismantled by presenting them as arbitrary and personal – curiously even if it is just to show the absurdity of such a position, thus undermining the very project of analysing the inscription of power into knowledge and vice-versa – and this is often legitimised through the rhetoric of critical evaluation of information as a personal responsibility and a matter of choice, consumer choice even. Once again, contemporary search engines, their presentation of results, and their flattening of everything into the same interface are clearly implicated in how this plays out, not least since using them is often called for in order to fact check claims circulating on social media. To spin further on the idea of a metaphorical marriage between deconstructing knowledge claims and critical evaluation of information, we suggest that other elements need to be wedded for it to work out. More precisely, this means that critical evaluation of information needs to be seen more strongly as a social venture and here social necessarily includes materiality and is in many cases entangled across various infrastructural arrangements, some of which are platforms acting according to the rules of corporate capitalism.

Let us now recycle the concept of friction of relevance as has been introduced earlier in the book. It seems as if we need to develop a critical preparedness to handle situations when the friction of relevance, that is, when the gap between relevance as constructed by search engines and the interests of society, is experienced as too large. In Europe in 2018, the so-called *Independent High-level Group on fake news and online disinformation* by the European Commission issued a report entitled "A multidimensional approach to disinformation" (European Commission 2018), where also promotion of media and information literacy is highlighted to "counter disinformation and help users navigate the digital media environment" (ibid., p. 5). Here media and information literacy is once again related to active citizenship, to empowerment, to critical thinking, and so on. What is stressed is the preventive function that media and information literacy should have to counter exploitation of the trust of citizens by more or less organised disinformation campaigns. What this report recognises, however, is the role that apart from commercial actors and foreign governments, domestic groups and even citizens play in actively undermining trust in political processes, the media and institutions (ibid., p. 11). In the report, this is not directly related to the promotion of media and information literacy in EU member states, but it does make visible a dissonance that is only rarely discussed.

The promotion of media and information literacy and, in fact, the very notion of media and information literacy is deeply permeated with a rational view of information and panders to ideals of an enlightened citizenship making well-informed choices, and it is furthermore entirely dependent on the functioning of the very institutions whose demotion it should protect against. While this is not necessarily a problem in all instances or concerning all types of knowledge, certainly not in a school context, it constitutes clearly a mismatch concerning the

much repeated mantra that media and information literacy should reinforce empowered citizenship in relation to democracy, its institutions and decision-making processes. Moving the question of media and information literacy entirely into the situation it emerges from equally fails to resolve this issue; rather it risks aggravating it. To put it bluntly, undemocratic practices and ideologies can be carried out by people who are highly media and information literate in relation to their respective communities and practices. It is easy to imagine, that disruptive activities, disinformation campaigns, or the sustenance of various antidemocratic conspiracy theories, are enacted by people acting within their social practices as very information literate.

There is often a mismatch between normative media and information literacy standards as promoted by international organisations and national agencies on one side and research starting from the premise of analysing specific situations on the other. Still, literacy in information literacy research is often treated as something progressive and largely positive. This is understandable; it is empowering and thus desirable to be able to obtain relevant information, to critically evaluate it, and to do this in relation to the specific conditions of a practice, a community and a situation. Often this does not clash on a fundamental level with the more prescriptive notion often tied to (liberal) democracy that is foregrounded in policy making and specifically in educational settings and sometimes, from a research-driven perspective, such a clash could be even productive. Cleary studying information literacy and information practices of groups with undemocratic, authoritarian, or illegal aims can be legitimate and certainly relevant. Yet, the question remains, what does this mean for how media and information literacy and critical evaluation of information are taught in educational institutions, when they are on the one hand positioned as nothing less than an unconditional prerequisite for democracy and enlightened citizenship, but on the other hand entirely submerged in just about any practice – at the same time as they are – in the reality of most people and most pupils – entirely dependent on ad-financed corporate surveillance tools beyond the control of the very democratic institutions their use is supposed to uphold?

Chapter conclusion: Locating trust and authority

At the core of this chapter, focusing on searching in relation to media and information literacy has been the absence of a contextual understanding of search as a content for education. By remediating different genres at the same time as they are converging and flattening those on the same interface, the curatorial functions of search engines become even more opaque and impenetrable. All this has implications for media and information literacy, as a policy concept, as a professional practice as well as in research. In order to address this, we propose the concept of infrastructural meaning-making; with this we want to focus the attention on the workings of infrastructure for search and other ways of accessing information, in order to put words on our take on information literacy and how knowledge is

dependent upon trust also in nonhuman actors. The ubiquity of search has far-reaching implications for didactics and what kind of knowledge is taught and valued in schools.

Developing media and information literacy – which is undeniably very important – can only partially address what is ultimately a crisis of trust in societal institutions of knowledge. In some respects, it might even exacerbate it. Someone opposed to vaccines will likely not evaluate vaccine-critical websites in different ways after a course on media and information literacy, whether contextually oriented or not. It is not possible to address a largely emotion-based understanding with intellectual arguments and critical evaluation of sources. Furthermore, being able to conduct a critical evaluation of information is nothing that those with the established knowledge base in society as a reference point have a monopoly on (compare Latour 2004). On the contrary, there are many examples of how critical evaluation of information as a method has evolved into an instrumental tool in the struggle between institutionally sanctioned and so-called alternative versions of knowledge, where both sides employ the same methods to reinforce their own position. As discussed in previous chapters, how we express what we search for, which sequence of words we use to convey to the search engines what we are looking for, is far from neutral. Rather it can be seen as enacting conformation bias. Often, people search not just for one or two terms, but insert a full sentence (Tripodi 2018), sometimes with a question mark at the end, and the way this sentence is phrased makes a big difference to how the search engine responds and calculates the relevance of the results it returns. Voice assistants together with natural language search are bound to further reinforce this way of engaging with search engines. In some sense, in the future, critical evaluation of information in all its variants must be discussed together with a discussion of how established knowledge and the institutions that produce, support and communicate it (libraries included) can maintain trust in knowledge institutions as well as in society at large. This remark is also relevant to the critical information literacy movement. Critical theory has no monopoly of the notion of being critical, and a "critical" perspective on media and information is nowadays also frequently delivered by anti-democratic forces. By referring to the limits of media and information literacy, it might sound as if we are contradicting ourselves, and in a way, we are, but this is a contradiction we as a society have to live with and perhaps it is being able to deal with this contradiction, this paradox, that constitutes critical information literacy.

References

Alexandersson, M. & Limberg, L. (2013). "Changing conditions for information use and learning in Swedish schools: A synthesis of research." *Human IT*, 11(2), 131–154.

Andrejevic, M. (2013). *Infoglut: How too much information is changing the way we think and now*. New York: Routledge.

Buckingham, D. (2006). "Defining digital literacy: What do young people need to know about digital media?" *Nordic Journal of Digital Literacy*, 1(2), 263–276.

Coughlan, S. (2017). "Schools should teach pupils how to spot 'fake news'." *BBC*. Retrieved from http://www.bbc.com/news/education-39272841.

D'Onfro, J. (2018). "Meet the man whose job it is to reassure people that Google search isn't evil." *CNBC*. Retrieved from https://www.cnbc.com/2018/06/22/danny-sullivan-on-being-google-new-search-liaison.html.

Elmborg, J. (2006). "Critical information literacy: Implications for instructional practice." *The journal of academic librarianship*, 32(2), 192–199.

European Commission (2018). *A multi-dimensional approach to disinformation: Report of the independent High level Group on fake news and online disinformation*. Luxembourg: Publications Office of the European Union.

Francke, H. & Sundin, O. (2009). "Format agnostics or format believers? How students in high school use genre to assess credibility." *Proceedings of the Association for Information Science and Technology*, 46(1), 1–7.

Gross, M. & Latham, D. (2012). "What's skill got to do with it? Information literacy skills and self-views of ability among first year college students." *Journal of the American Society for Information Science and Technology*, 63(3), 574–583.

Gärdén, C., Francke, H., Hampson, A., & Limberg, L. (2014). "A matter of facts? Linguistic tools in the context of information seeking and use in schools." *Information Research*, 19(4). Retrieved from http://InformationR.net/ir/19-4/isic/isic07.html.

Hardwig, J. (1991). "The role of trust in knowledge." *The Journal of Philosophy*, 88(12), 693–708.

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). "Trust online: Young adults' evaluation of web content. "*International journal of communication*, 4, 27.

Hellsten, I., Leydesdorff, L., & Wouters, P. (2006). "Multiple presents: How search engines rewrite the past." *New Media & Society*, 8(6), 901–924.

Hicks, A. (2013). "Cultural shifts: Putting critical information literacy into practice." *Communications in Information Literacy*, 7(1), 50–65.

Hillis, K., Petit, M., & Jarrett, K. (2013). *Google and the culture of search*. New York: Routledge.

Huvila, I. (2013). "In Web search we trust? Articulation of the cognitive authorities of Web searching." *Information Research*, 18(1) paper 567. Retrieved from http://InformationR.net/ir/18-1/paper567.htm.

Höchstötter, N. & Lewandowski, D. (2009). What users see: Structures in search engine results pages. *Information Science*, 179(12), 1796–1812.

Johansson, V. (2012). *A time and place for everything? Social visualisation tools and critical literacies*. The Swedish School of Library and Information Science: The University of Borås.

Julien, H. & Barker, S. (2009). "How high-school students find and evaluate scientific information: A basis for information literacy skills development." *Library & Information Science Research*, 31(1), 12–17.

Kammerer, Y. & Gerjets, P. (2014). "The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface." *International Journal of Human–Computer Interaction*, 30(3), 177–191.

Kammerer, Y. & Gerjets, P. (2012). "Effects of search interface and Internet-specific epistemic beliefs on source evaluations during web search for medical information: An eye-tracking study." *Behaviour & Information Technology*, 31(1), 83–97.

Kapitzke, C. (2003). "Information literacy: A positivist epistemology and a politics of outformation." *Educational Theory*, 53(1), 37–53.

Latour, B. (2004). "Why has critique run out of steam? From matters of fact to matters of concern." *Critical Inquiry*, 30, 225–248.

Lewandowski, D., Kerkmann, F., Rümmele, S., & Sünkler, S. (2018). "An empirical investigation on search engine ad disclosure." *Journal of the Association of Information Science and Technology*, 69(3), 420–437.

Limberg, L. (1999). "Three conceptions of information seeking and use." In. T. D. Wilson & D.K. Allen (Eds), *Exploring the contexts of information behaviour: proceedings of the second international conference on research in information needs, seeking and use in different contexts, 13–15 August 1998, Sheffield, UK* (pp. 116–132). London: Taylor Graham.

Limberg, L., Alexandersson, M., Lantz-Andersson, A., & Folkesson, L. (2008). "What matters? Shaping meaningful learning through teaching information literacy." *Libri*, 58(2), 82–91.

Limberg, L. & Folkesson, L. (2006). *Undervisning i informationssökning*. Borås: Valfrid.

Lloyd, A. (2014). "Following the red thread of information in information literacy research: Recovering local knowledge through interview to the double." *Library & Information Science Research*, 36(2), 99–105.

Livingstone, S., Van Couvering, E., & Thumin, N. (2008). "Converging traditions of research on media and information literacies." In J. Coiro, M. Knobel, C. Lankshear & D. J. Leu (Eds), *Handbook of research on new literacies* (pp. 103–132). New York: Lawrence Erlbaum Associates.

McNicol, S. (Ed.) (2016). *Critical literacy for information professionals*. London: Facet.

Mager, A. (2009). "Mediated health: Sociotechnical practices of providing and using online health information." *New Media & Society*, 11(7), 1123–1142.

Marchionini, G. (2006). "Exploratory search: From finding to understanding." *Communications of ACM*, 49(4), 41–46.

Metzger, M. J. (2007). "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research." *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091.

Metzger, M. J., Flanagin, A. J. & Medders, R. B. (2010). "Social and heuristic approaches to credibility evaluation online." *Journal of Communication*, 60(3), 413–439.

Mitchell, A., Jurkowitz, M., & Olmstead, K. (2014). *Social, search and direct: Pathways to digital news*. Washington D.C.: Pew Research Center.

Noble, U. S. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.

OECD (2018). "Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework." Retrieved from https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf.

Ofcom (2018). *Adults' media use and attitudes report*. London: Ofcom.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). "In Google we trust: Users' decisions on rank, position, and relevance." *Journal of Computer-Mediated Communication*, 12(3), 801–823.

Pilerot, O. & Lindberg, J. (2011). "The concept of information literacy in policy-making texts: An imperialistic project?" *Library Trends*, 60(2), 338–360.

Purcell, K., Brenner, J., & Rainie, L. (2012). *Search engine use 2012. Pew Research Center's Internet & American Life Project*. Washington, D.C.

Rieh, S. Y., Collins-Thompson, K., Hansen, P., & Lee, H. J. (2016). "Towards searching as a learning process: A review of current perspectives and future directions." *Journal of Information Science*, 42(1), 19–34.

Rieh, S. Y. (2002). "Judgment of information quality and cognitive authority in the Web." *Journal of the American Association for Information Science and Technology*, 53(2), 145–161.

Rosling, O. (2017). "Ola Rosling: Så kan vi uppfylla pappas dröm. "*Expressen*. Retrieved from https://www.expressen.se/debatt/ola-rosling-sa-kan-vi-uppfylla-pappas-drom/.

Rosling, H., Rosling Rönnlund, A., & Rosling, O. (2018). *Factfulness: Ten reasons we're wrong about the world – and why things are better than you think*. London, UK: Sceptre.

Savolainen, R. (2007). "Media credibility and cognitive authority: The case of seeking orienting information." *Information Research*. 12(3). Retrieved from http://www.informationr.net/ir/12-3/paper319.html.

Shirky, C. (2009). "A speculative post on the idea of algorithmic authority." Retrieved from http://www.shirky.com/weblog/2009/11/aspeculative-post-on-the-idea-of-algorithmic-authority/.

Simon, J. (2010). "The entanglement of trust and knowledge on the Web." *Ethics and Information Technology*, 12(4), 343–355.

Sundin, O. (2015). "Invisible search: Information literacy in the Swedish curriculum for compulsory schools." *Nordic Journal of Digital Literacy*, 10(04), 193–209.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices. "*Journal of Documentation*, 72, 990–1007.

Sundin, O. & Francke, H. (2009). "In search of credibility: Pupils' information practices in learning environments." *Information Research*, 14(4) paper 418. Retrieved from http://InformationR.net/ir/14-4/paper418.html.

Sundin, O., Haider, J., Andersson, C., Carlsson, H., & Kjellberg, S. (2017). "The search-ification of everyday life and the mundane-ification of search." *Journal of Documentation*, 73(2), 224–243.

Swedish National Agency for Education (2016). *IT-användning och IT-kompetens i skolan Skolverkets IT-uppföljning 2015*. Stockholm: Skolverket.

Taraborelli, D. (2008). "How the Web is changing the way we trust." In A. Briggle, K. Waelbers, P. A. E. Brey (Eds), *Current issues in computing and philosophy* (pp. 194–204). Amsterdam: IOS Press.

Tewell, E. (2015). "A decade of critical information literacy: A review of the literature." *Communications in Information Literacy*, 9(1), 24–43.

Tewell, E. (2016). "Toward the resistant reading of information: Google, resistant spectatorship, and critical information literacy." *Portal: Libraries and the Academy*, 16(2), 289–310.

Tripodi, F. (2018). "Searching for alternative facts: Analyzing scriptural inference in conservative news practices." *Data & Society*. Retrieved from https://datasociety.net/output/searching-for-alternative-facts/.

Tuominen, K., Savolainen, R., & Talja, S. (2005). "Information literacy as a sociotechnical practice." *The Library Quarterly*, 75(3), 329–345.

Wineburg, S., & McGrew, S. (2017). "Lateral reading: Reading less and learning more when evaluating digital information." *Stanford History Education Group Working Paper No. 2017*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3048994.

Wineburg, S., McGrew, S., Breakstone, J., & Ortega, T. (2016). "Evaluating information: The cornerstone of civic online reasoning." *Stanford Digital Repository*. Retrieved from https://ed.stanford.edu/news/stanford-researchers-find-students-have-trouble-judging-credibility-information-online.

UNESCO (2017). "Media and information literacy." Retrieved from http://www.unesco.org/new/en/communication-and-information/media-development/media-literacy/mil-as-composite-concept/.

UNESCO (2018). "Medie- och informationskunnighet: En demokratifråga." Swedish National Commission for UNESCO.

Wilson, P. (1983). *Second-hand knowledge: An inquiry into cognitive authority*. Westport, Conn.: Greenwood Press.

6

RESEARCHING SEARCH AND SEARCH AS RESEARCH

How do you as a researcher approach the invisible empirically? How does one get people to reflect on obvious but increasingly obscure everyday acts in a way that is both respectful and meaningful? The way we have approached this is by discussing search and search engines in focus groups. A focus group could be a weekly dinner party with a group of six friends gathering for a simple midweek meal; a group of researchers largely unknown to each other getting together over coffee at the university; teachers gathering during a lecture-free slot; university students working on a project setting aside some time for the discussion; or a group of pupils in a school setting. A researcher would function as a moderator and would present the project and its purpose and after making sure everyone understood their role and rights in it, the participants would be asked to anonymously write down the latest three occasions they remembered searching. Following this exercise, a focus group discussions could then for example start like this:

The researcher is sitting down with six 13-year olds for a focus group discussion about web searching. They have just written down the last three occasions when they have searched for something online. The youngsters know each other from school; in fact, the discussion is conducted at school even if the content of the discussion itself is not only or even primarily about school:

RESEARCHER: I thought we could start by talking about searching in a broad sense and ask whether you could tell us what you last searched for and how you went about it.

PERSON 1: First I searched on YouTube for a gaming video; Europa Universalis is the name of the game. Then I went to, then I was in my email, on Hotmail and I needed to find an Illuminati image.

PERSON 2 : Let's start. I am a very uninteresting person. I was on Facebook I guess, checked out the social network, Netflix (laughs). I can't remember which

DOI: 10.4324/9780429448546-6

movie I watched last, but I think it was Mr Nobody. Then I went onto Forever21 because I was looking for a shirt.

PERSON 3: Abucabar on Google for a writing project. Children who fight against child soldiers. And then Swefilms. And I don't remember which film I watched. And then Tv.nu [tv-listing]. There I looked for what I would see on TV.

[Someone in the background is heard saying that they have also been there]

PERSON 4: I have looked for like texts and quotes. I don't really know, I look for that very often. I look in Google then. Then I was on Swefilmer and watched a series. And I have looked for, I love food, I find it very entertaining to watch people prepare food and I have dedicated my life to watching people bake cakes on YouTube. Cakes (laughs) (more people join in laughing).

[Someone asks: "On YouTube?"]

PERSON 4: Yes. How to bake cakes and such like. Can be good to know

[Jokes about Gordon Ramsey and Kitchen Nightmares]

PERSON 5: Ok, I was on a shop and shop comedy /…/ and then I googled India.

[Laughter in the background]

PERSON 6: Yes and I googled "abubacar" on Google.com. And then I have googled properly on Google.com for heartstone download. A game.

The stories about search and similar web activities are told often without much involvement on the part of the researcher. Google and YouTube are mentioned, as well as certain web sites. The different activities the participants mention appear as so integrated with each other they are difficult to distinguish between. Facts are intermingled with entertainment, school work with gaming, cooking or watching movies, routine searches alternate with one-off questions. Google, which is, as in most of our focus groups the only search engine mentioned, brings about a context collapse between school and leisure – typical of search engines and social media. Search engines are part of everyday life, they co-construct it as their use contributes to interlacing its various temporal orderings.

Repeatedly in this book we have emphasised the ubiquity of online search in all walks of life. In this somewhat shorter chapter, we touch upon the issue of method. Having said that, the question of method is an immensely complex area and it is not our intention here to engage with it in detail, nor is the purpose of this chapter to provide a research method manual. Rather we want to broach some overarching reflections that are, on the one hand, tied to our own empirical material and methodological starting points and, on the other hand, arise from concerns characteristic for doing research on searching and search engines and of the particular areas of research we navigate.

The focus is on qualitative methods and to a lesser degree quantitative methods. This could be seen as rather old-fashioned at a time when data methods combining the two, or at least mixed methods, are often favoured – particularly if searching is investigated – an activity that so clearly leaves quantifiable digital traces. Steinar Kvale and Svend Brinkman (2009, p. xvii) introduce their book on interviews with, "[i]f you want to know how people understand their world and their lives,

why not talk with them?". That is basically what we have done in the focus groups that we refer to throughout the book – talked to people in various roles, of different ages and with different experiences. Or more precisely, we have, following the ideal of focus groups (Morgan 1996), listened to them talking with each other. This way we have encountered teenagers, university students, researchers, teachers, professionals, and so on. We have heard about their experiences of searching as well as not searching and we have discussed the technologies involved and platforms used. Except for the teenagers, most of our participants share the characteristic of being relatively well educated, which, of course, might have implications for what they talk about and how they reflect on the issues at stake.

The present chapter consists of two short parts approaching the issue from different directions. In part one, problems related to method when attempting to study processes that are largely invisible and often taken for granted are broached. Here we return to the two information science traditions of information retrieval and information behaviour, but add nuance by reflecting on how methodological choices are tied to specific epistemic objectives. The second part deals with how the use of search engines can be used for researching society and specifically informational forces at work in society. Here we touch upon the issue of digital methods and present selected approaches where the search engine is turned back onto itself or onto its surroundings.

How to investigate search and searching

Information retrieval and information behaviour research constitute the two largest research areas within information science, or library and information science (Togia & Malliari 2017). Although both investigate similar empirical phenomena, they tend to approach their study objects very differently. While information retrieval research is predominantly a quantitative endeavour, information behaviour research has diversified. In the early years it was largely survey-based and thus equally quantitative, yet the field has since methodologically branched out and embraced more and more qualitative methods, including various types of interviews, diaries, observation, focus groups, think aloud and so on (Bawden & Robinson 2015, p. 199). The literature is not unequivocal. However, it seems that a trend towards using qualitative research in information behaviour studies can be traced back to the 1990s (McKechnie et al. 2002) and a recent study shows a fairly even split between qualitative and quantitative approaches (Togia & Malliari 2017). This shift from a dominance of quantitative research towards an increasing use of qualitative research reflects the gradually deepening relation of information behaviour with the social sciences and humanities at the same time as information retrieval research has maintained a close relation with computer science. It should be noted here that we distinguish between quantitative and qualitative research respective quantitative and qualitative data collection methods. While the former is the starting point for the type of research questions asked, and thus the type of answers you can get, the latter are the concrete methods for collecting data.

Starting point in the user or the system?

At the end of the 1970s and beginning of the 1980s, proponents of what is often referred to as user-centred research in contrast to system-centred research became more and more visible in information science, as we also remark on in Chapter Two (e.g. Brenda Dervin, Nicolas Belkin, and Carol C. Kuhlthau). One text in particular has expedited this distinction, namely Brenda Dervin and Michael Nilan's paper "Information needs and uses" in the *Annual Review of Information Science and Technology* from 1986. With the paper a new "paradigm" of user-centred research was established. The text became a classic and has since then formed the research community's methodological self-understanding in many ways. The authors challenged the research community to change its starting point – instead of starting in a system and investigating how that has been used in quantitative terms, the research should start in the complex reality of users and from there on see what contextualised role information has in their lives. In methodological terms, Dervin and Nilan (1986, p. 16) argue:

> In the context of the traditional paradigm, quantitative approaches are seen as most compatible with traditional assumptions. Yet in context of the impetus of the paradigm shift, scholars are now calling for supplementing quantitative approaches with inductive, qualitative approaches.

While the claim of a paradigm shift might be overstated, it has led to a change in how research into either predominantly systems or users is perceived (Talja & Hartel 2007). Sanna Talja and Jenna Hartel conclude an investigation of user studies from the 1950s and 1960s by suggesting: "'systems-centredness' is easily associated with cold technology, indifference to individual human beings and their problems. 'User-centredness' easily appears as being ethically sound, warm and compassionate, based on the ethics of caring about people and their problems" (Talja & Hartel 2007). Could it be argued that the narrative of a schism between user and system-centred research as facing each other as two distinctly separate "paradigms" has led to diminished technological sensitiveness in some contemporary information behaviour research? Could it also be the case, as some longitudinal studies seem to indicate (Vakkari 2008; Julien, Pecoskie, & Reed 2011), that the growing share of qualitative methods in information behaviour research has resulted in more descriptive and theoretically less well-informed studies? Finally, is such a weakened interest in theoretically informed or explanatory research connected to an increased focus on the individual user, rather than on their positions in relation to various social contexts, as Vakkari (2008) has argued? We pose these questions as a means to direct attention to some of the methodological foundations upon which research on search, searching and systems for doing so is built and thus to dissect some of the assumptions built into different approaches, assumption which are rarely spelled out and thus typically go unnoticed.

Explanation and understanding

We want to complicate the picture by first introducing a distinction between *explanation* and *understanding* in order to subsequently consider different types of explanations. While information behaviour studies are often situated within the humanities and social sciences, calls for more explanatory research seem to spring from a position within the field closer to information retrieval with an understanding of theory (and models) that differs from the interpretative approaches of the humanities and social sciences, and also draws inspiration from psychology. Clearly, different understandings of the role of theory co-exist within information behaviour research (e.g. Fisher, Erdelez, & McKechnie 2005; Wilson 2013). This can be seen as indicative of the heterogeneity of the field and of research on search and search engines in particular. The late philosopher of science Georg Henrik von Wright (1971) sums up the different ways of approaching science. With references to, among others, Wilhelm Dilthey (1833–1911), who played an important role in the advancement of hermeneutics, he describes a split between explanation and understanding, where understanding has its roots in the humanities and explanation in the natural sciences. Von Wright (1971, pp. 6–7) further argues how the social and behavioural sciences, being in the middle of the two poles, constitute "a battleground for the two opposed trends in the philosophy of scientific method". Many of the information seeking and retrieval models, some of which we touch upon in Chapter Two, are motivated by an explanatory framework – that is, to present causal relations in order to explain why something has been done and to predict what will be done. When understanding is the goal, qualitative descriptions of fewer individual cases are a way to achieve that. When Donald Case and Lisa Given (2016, pp. 190–211), in their rich, almost overwhelming, overview of information behaviour research 35 years later, differentiate between an objectivist and an interpretative paradigm, it resembles the division between explanation and understanding. Information behaviour researchers closer to information retrieval research often strive for causal explanation, while much – albeit not all – research with a qualitative approach strives for contextualised understanding. Accordingly, Case and Given (ibid.) locate, for example, practice theory within the interpretative paradigm.

The notion of explanation is interesting, since it brings another far-reaching distinction to the fore, namely between *causal* explanation and *teleological* explanation. We want to use a short discussion from philosophy of science of this classical distinction to show the origins of some major differences in research approaches and how these are reflected in studies of information systems and information behaviours. In different epistemological settings, explanation can mean different things. Somewhat simplified, the natural sciences often seek causal explanations. They are interested in establishing causes that can be found somewhere else, typically prior to that which is being explained. In the social sciences and the humanities, on the other hand, another type of explanation is often sought – so-called teleological explanations. This type of explanation wants to understand the reasons,

intentions, or motivations for doing something. Teleological explanations can be more explicit and investigated as being held by individuals, or they can be captured as part of practices. Theodor R. Schatzki (2001, p. 47) also talks about affectivity, that is, "how things matter" to people, as an additional "factor" for understanding practices. Teleological explanations and affectivity together form, in his view, a "teleoaffective structure" (ibid., p. 50) that provides people with certain accepted ways of doing things and they are implicated in organising practices. Similarly, the sociomateriality of a practice makes certain types of agencies possible and any study of practices in which online searching is a part needs not just to investigate the teleological explanations of human participants but also how the various sociomaterial arrangements make them viable and certain outcomes more likely and desirable than others. As Wanda J. Orlikowski (2007, p. 1445) formulates it, foregrounding very much the materiality, digital and otherwise, of the search engine as an object summoned into various practices: a "Google search is constituted by the performativity of computers, networks, software, algorithms, directories, databases, and infrastructure, as these are enacted by the human agencies entailed in their design, construction, and operation".

"[T]he narrative of the user-centred turn", as Talja and Hartel (2007) identify, advances studies with a starting point in the user which notice information systems only as, and if, they are perceived by the user. In contrast, the so-called practice turn in the social sciences as it played out in different parts of information science moved the base from which to pose research questions from the user to the practices, and in most social practices the use of dedicated information systems is but one, very conditional and often deeply submerged, element. As a result, information systems such as web search engines seem to have taken an even less prominent place in information behaviour research. This has implications for our own work, as we also remark on in different parts of this book. How can we investigate peoples' dealings with a certain phenomenon such as search engines at the same time as we maintain searching must be investigated as entangled across social practices? Furthermore, if we argue that search has largely become invisible, particularly when the issue is its role in everyday life, how can we "capture" what we study? In this respect, a focus on information retrieval – either in a lab setting or "in the wild" – seems to be much easier for "capturing" searching as such. Besides, we also want to situate searching (and search engines) in a broad cultural, social, and political landscape. For this, an approach advocating the isolation of variables in order to explain their causal relations is not going to work. That does certainly not mean that research committed to establishing causality is not valid or useful. On the contrary, it evidently is, and we build on it wherever it is needed to advance our argument. We do, however, maintain that, if information science is to provide tools for understanding the role of search and search engines in everyday life, it must also be bold enough to paint a broader picture where empirical research together with theoretical insights can provide an understanding of not just what we do with search engines, but what search engines at the same time do with us and ultimately with society at large.

Seeing search in everyday life

Let us return to the question posed above: How can we go about investigating people's dealings with search engines at the same time as we maintain that searching must be investigated as entangled across social practices? As we showed in Chapter Two, earlier research on information seeking in everyday life, sometimes referred to as ELIS, rarely pays any attention to search or search engines. Could this have to do with the way in which the use of search engines is naturalised to such a degree that it does not surprise the researcher sufficiently and is thus neglected? Or might it have to do with the way in which, for instance, interview or observation guides steer the researcher to look beyond the boundary the system is seen to enact, as a relic from user studies and in a deliberate move to "unsee" the system? A focus on practices characteristically also means a focus on material objects, things that are involved in how something is typically done and which are, in turn, formed, also in their meaning, in relation to this doing. In his overview of practice theory in information science, Andrew Cox (2012) clearly shows the importance of taking account of material objects (see also Lloyd 2010; Pilerot 2014) also in information related studies. Yet, digital objects, such as search engines, are notoriously difficult in this regard. They are elusive, multi-layered and, as we also argued in Chapter Three, multiple. As such, they are packed into other technological objects, flattened on a screen or into a smart speaker, and, not to forget, they are often encased by proprietary code and subject to myriads of legal restrictions. Search engines, like other digital objects, and how people engage with their materiality, are in other words difficult to observe and this poses a significant methodological challenge. Information practice research has often drawn on methods inspired by ethnography, including data collection methods such as observation and less strictly structured interviews. The potential challenge mentioned above is clear. Search and search engines are changing our society as well as our ways of living in it, yet the number of searches carried out in one day by any one individual might still be quite few. How can we capture them through observation if people, for instance, only search occasionally and without any particular pattern? One way to circumvent this is the use of structured information diaries (Hektor 2001; McKenzie 2003; Rieh 2004), yet due to its ubiquity the particulars of a search might not even always be recognised as a noticeable activity. Moreover, when people do search, the activity is mostly hidden between a screen and a person as well as, of course, behind the screen. The challenges posed in researching voice search are even more complex (Sünkler et al. 2016), not least regarding questions of integrity, and we foresee an intensified method discussion as emerging in all areas of the field in relation to this development.

Regarding search and search engines, it seems as if there is a need to bring some aspects of their role into view by explicitly discussing them. Yet, general purpose search engines in particular are also hard to pinpoint by means of conversational methods, as it is difficult to strike a balance between zooming in and zooming out, getting too close and looking from too far away. The quotes from our own

research which we drew on in Chapter Four and to some extent in Chapter Five were collected through focus groups with more than 120 participants during the course of about one year.[1] All focus groups started with the moderator asking the participants to write down what they had recently searched for in order to focus the participants' thinking on the issue at hand. This was followed by a discussion. That is, we provided a lot of space for peoples' experiences and reflections about searching and search engines. At the same time, the agency of search engines in general and more specifically, for example, algorithms, data, indexes, and results pages are explored through earlier research, coming from both information retrieval and studies of the social shaping and politics of search and search engines.

The choice to conduct focus groups was made in order to let people reflect on activities that are most often invisible and integrated into their everyday lives and to have them do the reflecting together, which chimes well with a practice approach trying to elucidate the "typical" rather than the specific. In some of the focus groups, we also asked the participants to bring search devices (laptop, smartphone, tablet) to the discussions and occasionally they used them and commented on their use. However, this element posed a challenge, since, due to privacy concerns, we could not ask our participants to share screenshots with us, which likely would have enriched our material. It was also difficult to embed the use of these devices into the focus group discussions without interrupting the "natural" flow of the conversation and thus making search engines too strange and too abstract for our purposes.

Information retrieval

As already mentioned in Chapter Two, information retrieval is a more homogenous field of research than information behaviour research. Information retrieval research is in general motivated by the ethos of developing better information systems and its research does this primarily by, somewhat simplified, trying to model user behaviour and evaluate information systems in relation to relevance (Carterette, Kanoulas, & Yilmaz 2012, p. 106). This can, however, be achieved through different methodological approaches. Traditional information retrieval research assesses relevance in test collections. It does this in relation to measures based on the basic idea of distinguishing between precision and recall, where precision refers to a high degree of relevance in the returned documents, while recall refers to as many relevant documents as possible among the returned documents. The influential TREC conference and its test collection method, here a representative of traditional information retrieval, often assigns the user a background role (if any). Peter Ingwersen and Kalervo Järvelin (2005, pp. 4–6) distinguish between traditional information retrieval, which they refer to as the Laboratory Model, and interactive information retrieval in the following way:

> In this view [the Laboratory Model] real users and tasks are not seen as necessary. Test requests typically are well-defined topical requests with verbose descriptions that give the algorithms much more data to work with for query

> construction than typical real-life IR situations (e.g., web searching) do. Relevance is taken as topical, but factual features (based on structural data items, like author names and other bibliographic features) could be included. Relevance also is static, between a topical request and a document as seen by an assessor.

The spread of user studies and of interactive information retrieval studies of various kinds changed the approach to relevance from a relation between system and its documents (including their representation) to a relation between the user and a document. Ingwersen and Järvelin (2005, pp. 237, 389) distinguish higher order relevance from lower order relevance, where the lower level is considered to be relevance as a static object (topical relevance), as in traditional information retrieval, and higher order is regarded as different subjective types of relevance (but seen in relation to work task, situation, and so forth). A further quantitative approach within information retrieval to understanding the use of search engines is so-called log mining, that is the analysis of the automatically generated files logging the interactions of users with a search engine (Carterette, Kanoulas, and Yilmaz 2012), also called web log analysis (Jansen 2006).

The typical information retrieval researcher tries to come up with explanatory causal relations between variables. The specific data collection methods in information retrieval are numerous – everything from surveys and interviews to log analysis and online evaluation (Ingwersen & Järvelin 2005, pp. 191–256; White 2016, pp. 337–360). Even if a number of qualitative methods are also mentioned in Ryen White's (2016) overview, such as interviews, focus groups, and think aloud, it is safe to say that the majority of research is quantitative, where methods more common to qualitative research are used as complements or during the early, exploratory phases of a development process. Phenomena investigated include search query formulations, number of clicks, web trails, dwell time estimation and the movement of the gaze across the interface (eyetracking) (ibid., pp. 21–53). These are observed both in laboratory settings – that is in a controlled environment that is more or less constructed in order to mirror searching outside the laboratory – and in so-called natural settings, that is "in the wild" as the phrase often goes (ibid., pp. 54–57).

One out of many interesting examples where logged data are used that could methodologically be unpacked is Bernard J. Jansen, Danielle L. Booth and Amanda Spink's (2008) often-cited study on the classification of search queries through search logs. They automatically classified more than 1.5 million search queries from three different search engines (none of them Google) and showed how 80% of the log files they analysed were informational (finding content on a topic), while 10% were transactional (finding a website with the intent of purchasing a product) and 10% were navigational (finding a particular website). This classification of web searches according to the assumed intent of a query, i.e. which type of information need a query responds to, goes back to a much-cited paper by Andrei Broder (2002) from the early 2000s. In the study Broder combined a survey of Alta Vista users with an analysis of web logs obtained from the same search engine. The

validity of Broder's findings has since been contested (Lewandowski, Drechsler, & von Mach 2012), specifically by addressing the issue of method. Dirk Lewandowski, Jessica Drechsler and Sonja von Mach (2012) conducted a large-scale comparison of three studies each with a different approach to classifying query intent; a crowdsourcing approach, clickthrough data from search engine logs in combination with human raters, and an online survey. They conclude that existing methods for classifying query intent lack coherence. Thus, the resulting classifications of web searches according to the assumed intent of a query are problematic and, we suggest, point to a methodological challenge in trying to quantify human perception.

In Jansen, Booth, and Spink (2008) as well as Broder (2002), the researchers used web searches that were done in "real life", outside the lab. This brings up an issue with significant methodological implications. What can actually be concluded about web searches and contemporary general-purpose web search engines by studying information retrieval under entirely artificial conditions, and which different approaches are feasible? For instance, Dirk Lewandowski (2015) draws attention to problems with what he calls TREC-style settings to evaluate the retrieval effectiveness of web search engines, and also highlights that many of the "classic" studies are not only dated, but are also based on a small amount of data in comparison with the actual numbers of searches carried out in contemporary web search engines. One of the problems, however, in studying retrieval effectiveness and similar issues "in the wild" is, for one, the sheer enormity of real life search data and secondly also the lack of direct access to the data of commercial search engine providers. Google specifically does not allow for automated querying. Dirk Lewandowski and Sebastian Sünkler (2013) circumvent this, for instance, with a screen-scraping tool, which allows for compiling a larger, more controlled material, to then be assessed by jurors, which is, however, still more representative of real-life situations than a pure laboratory setting would render possible.

Eyetracking is a method that has advanced quite prominently in order to get to grips with how people interact with web search engines, as it enables researchers to see how the gaze moves across the screen prior to actually making a decision, as in a click on a link. Bing Pan and colleagues (2007) use a combination of eyetracking and click data in their important and much cited study on trust in Google's ranking, through letting a limited number of undergraduate students, which represents the typical population in most IR studies (Lewandowski 2015), complete ten search tasks – five navigational and five informational questions. This study was carried out in a lab, in a controlled environment and based on search queries developed by the research team rather than in everyday life. The significance and timeliness of the tasks for the participants were instead established through using a tool provided by Google to establish popular searches at any given time (Google Zeitgeist, a precursor to Google Trends). A proxy server made it possible to rearrange some of the results and thus manipulate Google's perceived ranking. One of the hypotheses tested concerned the likelihood of trusting Google, despite manipulation of the order of ranking.

> Thus, we anticipated dissociation between the ocular data, which would indicate some implicit conflict between the position and the actual Google rank, yet that the subjects would still choose a higher positioned abstract based on a greater trust in Google's algorithm than in their own judgement.
>
> *Pan et al. 2007, p. 812*

Through detailed and quantitative analysis of where the participants looked on the screen (eyetracking), of what links they clicked on and by measuring the time spent on each task, they were able to confirm their hypothesis.

Using search engines for investigations

Search engines are not only objects of research – they can also be used as tools to understand social phenomena beyond searching itself, to zoom in on some of "the informational dynamics" (Trevisan 2014) at play in society and be thus turned into "research machines" (Rogers 2013, p. 99). Relatedly, Filippo Trevisan (2014) maintains that "scholars have been more interested in talking about the role played by search engines in shaping society rather than exploring ways in which these tools can augment our knowledge of social, political, and economic trends". Thus, as the famous dictum of Google as a "database of intentions" (Battelle 2011) alludes to, search engines – and mostly Google – become a source of information also in a different sense; for example, what kind of search phrases are the most popular in different countries, which changes do these undergo over time, or what kind of authority does Google provide to certain issues through its ranking or autosuggest function, and what does that authority tell us about what is regarded as important in society. As Wiesław Pietruszkiewics (2012, p. 80) puts it: "The statistics gathered from the search engines usage reveal the interests, tendencies, and moods within society because searching for particular products, events, or terms occurs when such information is required".

Perhaps the most common tool for doing research with Google is Google Trends. Originally a marketing tool, Google Trends facilitates finding metrics about the frequency of search terms used from 2004 onwards or any time period after 2004. The data can also be filtered through, for example, country of origin, types of Google service used (news, web, etc.) and some other parameters. The general idea is that users' search behaviour tallies with what happens in a social reality, also outside the internet. Trevisan (2014) sees "[a]dvantages over traditional methods concerned with the identification of broad socio-political trends (e.g. public opinion surveys)" and he also notices "[a]ugmented opportunities to study the relationship between off-line events and online behavior, especially in conjunction with crisis". The most famous example of how search frequency analysis, which is what Google Trends makes possible, was used to anticipate a crisis is GoogleFlu Trends (and Google Dengue Trends) (Carneiro & Mylonakis 2009), where certain search terms were used as indicators for influenza spreading across society. This was, however, done by Google itself and, after its initial success in

2008, it has been discontinued because the model proved to deliver inaccurate predictions. Google has been accused of "big data hubris" (Lazer et al. 2014), not least for not adequately grasping the situatedness of the data feeding into the model.

If we return to the distinction between explanation and understanding touched upon above, the approach of Google FluTrends is motivated by the former, specifically by establishing causal explanations. However, within an interpretative paradigm directed towards understanding, Google Trends and similar tools can play an important role when it is put to work together with other methods. Also, Google Search itself and the traces it uncovers, such as related searches, autocomplete suggestions, and also search results and their order, can work in such configurations. Google Trends has, for instance, been used to analyse the interplay of search and legacy media in relation to an election (Trevisan et al. 2018), to predict changing tourist patterns (Jackman & Naitram 2015) as well as to investigate environmental issues, such as the recycling (Haider 2016), climate change awareness in different countries (Archibald & Butt 2018) or the public acceptance of resource efficiency policies (Bicket & Vanner 2016), to name just a few examples. In all these cases, Google Trends and the search frequency analysis it offers is but one element in assemblages where different methodological approaches are brought together and where it is assigned larger or smaller roles depending on the research design.

Google Search traces and Google Trends can be used as a thermometer not just for gauging public opinions and emotions (as a complement to, for example, surveys) in themselves, but also for how these opinions and emotions can be analysed in order to understand very disparate issues such as the movement of people, everyday practices, or how a societal controversy plays out (Eklöf & Mager 2013), and so on. Importantly, this needs to be undertaken with a thorough understanding of how search engines work, of their fluidity and of their role in feeding back into the very issues they portray. Acute awareness of the limits of the specific tool and of the limits of Google itself for the purpose of research is crucial. The most important limits are connected to the proprietary status of search engines and the resulting lack of control over which data is collected, in which ways and, of course, the fact that the data cannot be reproduced.

We would like to dig into two studies from information science to see what kinds of questions are asked and how Google trends and other digital applications can be used. Jutta Haider (2016) – one of the authors of this book – used, among other tools, Google Trends to investigate how waste sorting is constructed as a problem and by extension how information is structured through search, and Isto Huvila (2013) carried out a webometric study on the very topic of "search" in order to understand the cognitive authority of web searching. Haider's (2016, p. 390) article is introduced as following:

> In December 2013 Google released a list showing that during that year the third most popular search in Sweden in the category "how to" had been "how

> to sort waste". A year later, in December 2014, the sorting of waste featured again in the results that Google presented for searches carried out in Sweden. Yet, this time it was in the form of the third most popular search in the category "why" – "why to sort waste".

The metrics by Google Trends set the scene of the issue under investigation, waste sorting and recycling in everyday life in Sweden. Combinations of data creation methods were thereafter used to compile a diverse material: searching for the most popular phrases in Google in relation to waste, following suggestions of related searches, downloading keywords associated with waste sorting, and establishing different types of link relations of the top search results in order to roughly sketch which key actors and topics are involved in shaping the issue. This is combined with qualitative content analysis on different levels of the material. The important point for us here is that this study does not primarily aim at saying something about web searching or Google as such. Rather, searching (and Google) is used to create an understanding of waste sorting and recycling in Sweden in relation to the online information landscape it is situated within by means of seeing it through Google.

Huvila's (2013) article has an aim closer to the one in this book. The study aims at investigating "perceived cognitive authority and credibility of the act of searching information". He collected 805 so-called utterances about search and searching through webometric analyst software:

> The material for the present study consists of 805 utterances related to information searching and search engine use collected from the Web. The heuristically generated phrases used in the harvesting of the utterances are listed in Table 1. The heuristics was based on an in-depth exploration of Web-based discussion forums and blogs to discern the patterns of expressing failed and successful searching. Each of the chosen expressions was tested by using a Google search (http://www.google.com) and overviewing the first ten results for their relevance in the present study i.e., whether the utterances were related to searching or not. Different wordings (e.g., I searched in Internet) were tested and the final selection of utterances was based on the number of retrieved hits (phrases with a large number of hits were preferred) and their relevance to the topic (phrases with a low number of actual Web searching-related hits were omitted) in the Google search test.

Also here the results obtained by way of actually using the object of study, the search engine, were developed further through qualitative textual analyses. The study shows, among other things, the relatively high cognitive authority that people attribute to searching and search engines, thus supporting other research. The issue at stake is searching and search engines, but the method used is search itself, rather than asking people about search.

Chapter conclusion: Search engines as objects of and tools for research

The task of researching search and search engines joins researchers coming from many corners of academia. Computer science, information retrieval, and some parts of information behaviour research formulate their research questions most often with a view to explanation, while other parts of the field strive more for understanding (compare von Wright 1971) and teleological explanation. Undeniably, such a distinction is strongly simplified, but it elucidates some of the difficulties in constructing a coherent research paradigm in relation to search, which was evident already in Chapter Two. We see how information science researchers most often go about studying searching (or seeking) as a phenomenon in itself (or in relation to tasks of various kinds), while researchers from many other social sciences that became interested in online searching do this most often with a particular brand in focus – and mostly this is Google. While a focus on searching as a generalised activity strives at revealing patterns of use or at developing models and categorisations that could explain and predict how people use search engines, the focus on Google in all its variants is of a more idiographic nature – that is, to focus on one deeply contextualised case – to understand this giant and how it affects society. For example, when Hillis, Petit, and Jarrett (2013, p. 7) set out to understand how Google has reached its "socially consecrated status", it is a very different study from when Jansen, Booth, and Spink (2008) analyse web logs with queries from three different search engines in order to categorise them as either informational, navigational or transactional (Jansen, Booth, & Spink 2008). Both investigate search engines, but where Hillis, Petit, and Jarrett deliver a largely theory driven study of Google in society, Jansen, Booth, and Spink provide us with generic knowledge of web search behaviour. Both studies are very interesting, and useful for understanding the ubiquity of search in everyday life, but they are also entirely different. This brings us full circle back to our starting point, positing that search and the use of general purpose search engines are entwined in everyday life and society at all levels. The question of method is a huge, complex issue, and we can necessarily only scratch the surface. Studying search engines means studying society. This is true for studies focusing on the power of Google's algorithms as much as for those invested in evaluating relevance assessments, and taking the challenges and tensions arising from this methodological span seriously is paramount.

Note

1 The material collection has followed the ethical guidelines of the Swedish Research Council. The participants were informed in advance about the project and we have kept the anonymity of all participants. Participants under 18 years had informed consent forms signed by their guardians prior to their participation. These focus groups with teenagers were held by Cecilia Andersson and they are also part of the empirical material of her PhD project. The conversations were transcribed verbatim, but with adjustments to account for some of the irregularities of the spoken language. Most groups were carried out in Swedish and quotes from these conversations were translated into English by the

authors. This adds an additional layer of anonymisation. In this book the quotes are used as examples to bounce off discussions. For a more detailed description of method, analysis, results and ethical considerations we ask the reader to study our articles and those of our colleagues (Andersson 2017a, 2017b; Carlsson & Sundin 2017; Haider 2017; Kjellberg & Haider 2018; Sundin et al. 2017; Sundin & Carlsson 2016).

References

Andersson, C. (2017a). "The front and backstage: Pupils' information activities in secondary school." *Information Research*, 22(1), CoLIS paper 1604. Retrieved from http://InformationR.net/ir/22-1/colis/colis1604.html.

Andersson, C. (2017b). "'Google is not fun": An investigation of how Swedish teenagers frame online searching." *Journal of Documentation*, 73(6), 1244–1260.

Archibald, C. L., & Butt, N. (2018). "Using Google search data to inform global climate change adaptation policy." *Climatic Change*, 150(3–4), 447–456.

Battelle, J. (2011). *The search: How Google and its rivals rewrote the rules of business and transformed our culture*. London: Hachette UK.

Bawden, D., & Robinson, L. (2015). *Introduction to information science*. London: Facet Publishing.

Bicket, M. & Vanner, R. (2016). "Designing policy mixes for resource efficiency: The role of public acceptability." *Sustainability*, 8(4), 366. Retrieved from https://www.mdpi.com/2071-1050/8/4/366/htm.

Broder, A. (2002). "A taxonomy of web search." *ACM Sigir Forum*, 36(2), 3–10.

Carlsson, H. & Sundin, O. (2017). "Searching for delegated knowledge in elementary schools." In Proceedings of ISIC, *Information Behaviour Conference*, *Zadar, Croatia*, *20–23 September, 2016*: *Part 2. Information Research*, 22(1), paper isic1618. Retrieved from http://InformationR.net/ir/22-1/isic/isic1618.html.

Carneiro, H. A. & Mylonakis, E. (2009). "Google trends: A web-based tool for real-time surveillance of disease outbreaks." *Clinical infectious diseases*, 49(10), 1557–1564.

Carterette, B., Kanoulas, E., & Yilmaz, E. (2012). "Evaluating web retrieval effectiveness." In D. Lewandowski (Ed.), *Web search engine research* (pp. 105–138). Bingley: Emerald.

Case, D. & Given, L. (2016). *Looking for information: A survey of research on information seeking, needs, and behavior* (4th ed.). Bingley: Emerald.

Cox, A. M. (2012). "An exploration of the practice approach and its place in information science." *Journal of Information Science*, 38(2), 176–188.

Dervin, B. & Nilan, M. (1986). "Information needs and uses." In M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (pp. 3–33), 21. White Plains, NY: Knowledge Industry Publications.

Eklöf, J. & Mager, A. (2013). "Technoscientific promotion and biofuel policy: How the press and search engines stage the biofuel controversy." *Media, Culture and Society*, 35(4), 454–471.

Fisher, K. E., Erdelez, S., & McKechnie, L. (Eds). (2005). *Theories of information behavior*. Medford, N. J.: Information Today.

Haider, J. (2016). "The structuring of information through search: Sorting waste with Google." *Aslib Journal of Information Management*, 68(4), 390–406.

Haider, J. (2017). "Controlling the urge to search: Studying the informational texture of practices by exploring the missing element." *Information Research*, 22(1), CoLIS paper 1613. Retrieved from http://InformationR.net/ir/22-1/colis/colis1613.html.

Hektor, A. (2001). *What's the use? Internet and information behavior in everyday life*. Linköping, Tema: Linköping University.

Hillis, K., Petit, M., & Jarrett, K. (2013). *Google and the culture of search*. New York: Routledge.

Huvila, I. (2013). "In Web search we trust? Articulation of the cognitive authorities of Web searching." *Information Research*, 18(1) paper 567. Retrieved from http://InformationR.net/ir/18-1/paper567.html.

Ingwersen, P. & Järvelin, K. (2005). *The turn: Integration of information seeking and retrieval in context*. Dortrecht: Springer.

Jackman, M. & Naitram, S. (2015). "Research note: Nowcasting tourist arrivals in Barbados – just Google it!" *Tourism Economics*, 21(6), 1309–1313.

Jansen, B. J. (2006). "Search log analysis: What it is, what's been done, how to do it." *Library & information science research*, 28(3), 407–432.

Jansen, B. J., Booth, D. L., & Spink, A. (2008). "Determining the informational, navigational, and transactional intent of web queries." *Information Processing & Management*, 44(3), 1251–1266.

Julien, H., Pecoskie, J. J., & Reed, K. (2011). "Trends in information behavior research, 1999–2008: A content analysis." *Library & Information Science Research*, 33(1), 19–24.

Kjellberg, S. & Haider, J. (2018). "Researchers' online visibility: Tensions of visibility, trust and reputation." *Online Information Review*.

Kvale, S. & Brinkman, S. (2009). *InterViews: Learning the craft of qualitative research interviewing*. (2nd ed.). Los Angeles: Sage Publications.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). "The parable of Google Flu: Traps in big data analysis." *Science*, 343(6176), 1203–1205.

Lewandowski, D. (2015). "Evaluating the retrieval effectiveness of web search engines using a representative query sample." *Journal of the Association for Information Science and Technology*, 66(9), 1763–1775.

Lewandowski, D. & Sünkler, S. (2013). "Designing search engine retrieval effectiveness tests with RAT." *Information Services & Use*, 33(1), 53–59.

Lewandowski, D., Drechsler, J., & von Mach, S.(2012). "Deriving query intents from web search engine queries." *Journal of the Association for Information Science and Technology*, 63(9), 1773–1788.

Lloyd, A. (2010). "Corporeality and practice theory: Exploring emerging research agendas for information literacy." *Information Research*, 15(3) colis794. Retrieved from http://InformationR.net/ir/15-3/colis7/colis704.html.

McKechnie, L. E., Baker, L., Greenwood, M., & Julien, H. (2002). "Research method trends in human information literature." *New Review of Information Behaviour Research*, 3, 113–125.

McKenzie, P. J. (2003). "A model of information practices in accounts of everyday-life information seeking." *Journal of Documentation*, 59(1), 19–40.

Morgan, D. L. (1996). "Focus groups." *Annual review of Sociology*, 22(1), 129–152.

Orlikowski, W. J. (2007). "Sociomaterial practices: Exploring technology at work." *Organization Studies*, 28(9), 1435–1448.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). "In Google we trust: Users' decisions on rank, position, and relevance." *Journal of Computer-Mediated Communication*, 12(3), 801–823.

Pietruszkiewicz, W. (2012). "The computational analysis of web search statistics in the intelligent framework supporting decision making." In D. Lewandowski (Ed.), *Web search engine research* (pp. 79–102). Bingley, U.K.: Emerald.

Pilerot, O. (2014). "Making design researchers' information sharing visible through material objects." *Journal of the Association for Information Science and Technology*, 65(10), 2006–2016.

Rieh, S. Y. (2004). "On the web at home: Information seeking and web searching in the home environment." *Journal of the American Society for Information Science and Technology*, 55(8), 743–753.

Rogers, R. (2013). *Digital methods*. Cambridge, MA.: The MIT Press.

Schatzki, T. R. (2001). "Practice mind-ed orders." In T. R. Schatzki, K. Knorr-Cetina, & E. von Savigny (Eds), *The Practice Turn in Contemporary Theory* (pp. 42–55). London: Routledge.

Sundin, O. & Carlsson, H. (2016). "Outsourcing trust to the information infrastructure in schools: How search engines order knowledge in education practices." *Journal of Documentation*, 72(6), 990–1007.

Sundin, O., Haider, J., Andersson, C., Carlsson, H., & Kjellberg, S. (2017). "The search-ification of everyday life and the mundane-ification of search." *Journal of Documentation*, 73(2), 224–243.

Sünkler, S., Kerkmann, F., Quirmbach, S., & Schultheiß, S. (2016). "Sprachgesteuerte Websuche im Test: Ein Erfahrungsbericht aus der Perspektive von Auftraggeber, Auftragnehmer und Durchführendem." In C. Mayr & M. Pinzger (Eds), *INFORMATIK 2016* (pp. 211–225). Bonn: Gesellschaft für Informatik.

Talja, S. & Hartel, J. (2007). "Revisiting the user-centred turn in information science research: An intellectual history perspective." *Information Research*, 12(4) paper colis04. Retrieved from http://InformationR.net/ir/12-4/colis/colis04.html.

Trevisan, F. (2014). "Search engines: From social science objects to academic inquiry tools." *First Monday*, 19(11), 11–3 November 2014. Retrieved from http://www.firstmonday.dk/ojs/index.php/fm/article/view/5237/4157.

Trevisan, F., Hoskins, A., Oates, S., & Mahlouly, D. (2018). "The Google voter: Search engines and elections in the new media ecology." *Information, Communication & Society*, 21(1), 111–128.

Togia, A. & Malliari, A. (2017). "Research methods in library and information science." In S. Oflazoglu (Ed.), *Qualitative versus quantitative research* (pp. 43–64). IntechOpen. Retrieved from https://www.intechopen.com/books/qualitative-versus-quantitative-research/research-methods-in-library-and-information-science.

Vakkari, P. (2008). "Trends and approaches in information behaviour research." *Information Research*, 13(4) paper 361. Retrieved from http://InformationR.net/ir/13-4/paper361.html.

Von Wright, G. H. (1971). *Explanation and understanding*. London: Routledge & Kegan.

White, R. W. (2016). *Interactions with search systems*. Cambridge: Cambridge University Press.

Wilson, T. D. (Ed.) (2013). *Theory in information behaviour research*. Kallered: Eiconics Ltd.

7

CONCLUSION

This book opened with the simple sentence: "This is a book about search". In the chapters that followed, we unpacked this statement and showed how search, search engines, their increasing invisibility and entanglement in the practices of everyday life can be understood using tools from different traditions within the field of information science and from other fields of research. We also showed various relations that online search enters into with other elements in social practices and how search engines and their use are in different ways part of everyday life and are thus implicated in how we know, when we know, and how we organise knowing. We also discussed how search engines contribute to *factualising* things. In this process, we brought media and information literacy studies on board to help us discuss the difficulties involved in turning search and search engines into objects of learning and teaching in educational settings. Here, we also discussed the limits of common perceptions of what media and information literacy skills are expected to achieve in the contemporary media and information landscape.

The vast majority of the research on search in information science comes either from the field of information retrieval, with a focus on the retrieval process, or from the area of information behaviour, including information literacy, with its broader interest in the users and their practices. We argue that both these traditions provide important insights helping us to understand online searching in everyday life. Accordingly, we drew on insights from both so as to develop an understanding of search and search engines in everyday life. Yet, we added nuance by embedding this in a framework sensitive to how the materiality of practices embeds search engines into the complex sociality of everyday life and also an awareness of the socioeconomic, political, and cultural conditions implicated in shaping contemporary web search engines into platformised information infrastructures driven largely by corporate interests. The rest of the chapter is organised as follows: we begin with a short reflection on temporalities of search, a discussion which relates

DOI: 10.4324/9780429448546-7

back to the notion of everyday life as temporal introduced in Chapter One and which helps to lay additional emphasis on the profound implications of search engines for everyday life and its coordination. This is followed by briefly recapitulating some of the notions that emerge as central throughout the book. Finally, we conclude with a reflection on the role and also the stewardship of information science in understanding online search and search engines in society.

Temporalities of search

The different constitutive entanglements of social practices and information retrieval technologies, including web search engines, form part of an algorithmisation of everyday life (Willson 2017), which is likely to increase in the future and in many ways this algorithmisation is temporal. It concerns how information can be found online rather than being remembered, regarding the temporal variability of search results facilitated by different forms of personalisation, or how search engines continuously remake their own archive, thus engaging in memory politics (Zavadski & Toepfl 2019), changing the records of the past and society's access to it as well as the relationship between past, present and future (Chun 2008; Hellsten, Leydesdorff. & Wouters 2006). It is no coincidence that one of the most debated areas of search engine control concerns the so-called "right to be forgotten" by the search engines, directly pertaining to their involvement in how public memory is shaped in relation to the private sphere (Lindsköld 2018). The algorithmisation of everyday life through search engines also takes place in relation to the much commented on professed acceleration of society, as it not least plays out in the flexibilisation of the time regimes coordinating everyday life (Rosa 2013, p. 235) and in which search engines are complicit. At this point it is also interesting to speculate in terms of how the possibilities of predictive search constitute a form of algorithmic future-making thus further increasing the perceived acceleration of the rhythm of life that characterises late modernity. Much as postulated in the Thomas theorem, which says "[if] men define situations as real, they are real in their consequences" (Thomas & Thomas 1928, p. 572), here if a search engine says something should be searched for, it becomes part of your future, if the algorithm imagines it, it becomes real in its consequences.

Thus, what we want to communicate here is how the algorithmisation of everyday life also constitutes a number of temporal changes, in the sense that routines are arranged differently, thus re-structuring the rhythm of life. Time and temporality are not fixed, nor is time exclusively physical, abstract or measurable by clocks, rather, as not least the foundational work of Barbara Adam (1995, 2006) demonstrates at length, time is acutely social; it is multi-faceted, involved in social conventions and the "concrete relations between people" (Adam 1995, p. 20). Search engines as algorithmic intermediaries between people and information traverse our various engagements with and our making of time. The algorithm becomes, in the words of Mark Andrejevic (2013, p. 144),

> an organizing social principle: one in which the rhythms of our daily lives create patterns beyond our comprehension and our predictive abilities (but not that of the database) – patterns that are turned back upon us for the purposes of sorting, exclusion, management, and modulation.

Social practices change, slowly, but with significant implications for the entwinement of search engines as information infrastructures, across the practices of everyday life. Reijo Savolainen (2006, p. 119) reminds us, that "[t]ime is one of the main contextual factors of information seeking" (see also Hartel 2010), while Elizabeth Davies and Pamela McKenzie (2002, p. 4) maintain: "*When* an individual seeks information is as important as *what* he or she seeks". If we bring this reasoning together with Chuck Moran's (2015) notion of how practices with significant technological components organise temporality that we discussed in Chapter One, this opens up for an interesting sociomaterially informed understanding of the relationship between different notions of time and of search engines as tools for information seeking. Clearly, we could see that use of search engines as a technical component in social practices is far from inconsequential, but in profound ways coordinates the practices it is part of. Search engines are thus implicated in organising the temporality of everyday life.

The use of search engines occurs, as we have discussed at length throughout this book, embedded into networks of other activities, also information-related activities. These relations are partially dynamic, that is, while stable and recognisable in certain respects, they are also continuously emerging. This becomes most evident in accounts of how search engines are implicated in organising social relations and ways of knowing, when they are reflected on by means of notions of time. In this sense, time functions as a "discursive resource" (Davies & McKenzie 2002, p. 7) for structuring the accounts of how the embedding of search into everyday life is experienced, and for expressing some of the different roles that search engines play in social practices. This occurs, for instance, in relation to how conversations can be cut short by means of a web search, how restlessness creeps in, or the anticipation of speedy delivery of answers changes who is asked or not asked for information, how search results are evaluated, and in what way queries are formulated. It also goes hand in hand with a late modern time regime, which as Hartmut Rosa (2013, p. 235) emphasises, is more strongly coordinated by deadlines rather than organised by means of pre-defined time slots. Interestingly, this is also evident, perhaps unexpectedly, in project work in school settings where the "deadline", as we could see, is explicitly part of how search engines are put to use. Likewise, the possibility of literally walking around everywhere with a search engine in one's pocket or having it accessible by simply speaking up in a room with a smart speaker has implications for what is considered a routine and thus for how practices are organised, in space and in time. We could see examples of this throughout the book. When being ill or being a tourist means using a search engine, this makes Google (or Bing or Baidu) part of the practices enacted, but it necessarily also means that they are implicated in the temporal coordination of the various additional elements that are brought in.

Central notions

Search-ification and mundane-ification of search

Due to the ubiquity of search and search engines in parallel with their invisibility, online searching has melted into myriads of social practices. Information seeking is most often a means to an end rather than the end in itself. We are gardening, preparing dinner, doing homework, taking care of our health, and searching is often enmeshed in these practices. Rather than talking about search as an information practice in its own right, we talk about search as entangled across social practices. We can talk of a search-ification of everyday life that has occurred alongside a mundane-ification of search. This describes how the use of information retrieval systems has developed from an advanced skill that was part of the professional practices of information professionals into becoming an unnoticed element of everyone's different everyday life practices. Search engines are key actors in the fabric of everyday life and its informational texture, and they are increasingly invisible. They are invisible in terms of how they are implicated in practices and in coordinating the rhythm of everyday life, but also with regards to how they work. What assumptions about relevance order their results in a certain way, what decisions make specific terms relations possible, how is the index structured, or what data is brought together on which grounds and what is not? All these and similar issues are largely invisible to most people. Yet they are increasingly consequential for our understanding of the world and our place in it.

Searching for content and searching in practice

In relation to this invisible mundanity of online search, we have discussed two different ways of approaching searching that have informed information science research: search for content versus search as practice. When approaching search with content in focus, it becomes a question of getting access to information in terms of epistemic content. Such an approach takes its starting point in what a webpage (or similar document) contains, their *aboutness*. With a practice approach, the attention is rather focused on questions such as how search and search engines establish what is important within an area of practice. The focus shifts to how information provided through the search engine is given meaning in different situations. Data, users, algorithms, indexes, and so forth shape what there is to know and this is made sense of in and through concrete doings. We have argued that the two perspectives should be combined: online searching is something that both needs to be understood as conveying content and as happening in and through social practices. This is necessary not only in order to be able to fully appreciate how media and information literacy in relation to search engines can be conceptualised as concerning the specific practice and situation it is part of, but also as concurrently pertaining to broader societal conditions, including struggles and constraints.

Infrastructural breakdown and frictions of relevance

In a sociomaterially informed understanding of information infrastructure, the concept comprises technical dimensions as well as practices. This means that an infrastructure becomes an infrastructure through its use. Thus, the notion that infrastructures become visible on breakdown necessarily also includes breakdown and malfunctioning at the level of use. This includes situations when searching is not possible or not appropriate, when a certain search engine (typically Google) is not available, when search results conflict with certain expectations or when the autocomplete function suggests inappropriate terms. The latter two examples can also be understood in terms of frictions of relevance. The notion of frictions of relevance in the case of general-purpose search engines describes the experience of dissonance between different individual needs, societal interests and the vested interests of the stakeholders catered to by the business model of the multisided platform structure.

Infrastructural meaning-making and the limits of critical evaluation of information

Keeping true to the notion of information infrastructures as being constituted in and through use also means that critical evaluation of information needs to account for the infrastructural arrangements that give rise to and enable certain types of information. This in turn needs to be related to how they are shaped by their use. We have suggested calling such an engagement with information's specific infrastructural affordances infrastructural meaning-making. People tend to see search engines as neutral ahistorical tools and as politically and ethically unproblematic, if they notice them at all. Their relation to card catalogues, encyclopaedias, libraries, bibliographies, archives, and so on is opaque to most people. Infrastructural meaning-making is a way to move beyond critical evaluation of information at a content level to include also the reasons why certain information surfaces at all and that it does this in particular ways. It points to ways in which infrastructural conditions enable information to exist in a certain way in the first place and thus locates sense-making and consequently also criticism even at this level. At the same time, a growing awareness of the infrastructural conditions for information is not enough. Such a call for evaluation of all kinds of knowledge claims also runs the risk of leading into a dead end of critical argumentation. Can more criticism resolve a crisis of trust? Our answer is most likely, no, it cannot. Yet, since information is constitutively entangled with the various infrastructures, and most often these are corporate, multi-sided platforms, we must also consider their functioning in relation to how trust is undermined and built.

Final thoughts

Throughout the book, we have entered into dialogue with very different types of research, attempting to read together work from vastly different corners of academia and from our own discipline, thus bridging a wide methodological and theoretical span. Necessarily some of these exchanges were deeper, while others have remained on a

surface level. Nevertheless, we argue that these, and similar, conversations are needed in order for information science to pull its weight in efforts – within and outside of academia – to elucidate the ways in which today's corporate information infrastructure is implicated in shaping the informational texture of society and of how the increasing ubiquity and evolving invisibility of the digital transform everyday life, education and society at large. Relevance, recall and precision, the index, classification systems, information needs, citation indexing, and so on, all these are notions and concepts with a long history in information science, and library and information science, and more recently this tradition also includes critical perspectives on these very notions. These concepts are some of the most fundamental building blocks of contemporary web search engines, and the conversation alluded to above needs insights from the field that coined, developed and fought over them for decades.

At the beginning of this book, we introduced Google as the elephant in the room. It is by now clear that this analogy mostly holds true in relation to information science and in particular to work on information practices. Yet, having said that, maybe this analogy does not really capture what is at stake in the first place. Perhaps Google is better likened to the room itself. Mostly we do not notice it; we move in it and through it without thinking; we rarely reflect on the existence of walls or why the furniture is placed in a certain way. Yet occasionally something changes, a piece of furniture is moved or a picture is taken down and we notice the room and all its limits.

This book covers an inter-disciplinary field of research that is constantly expanding and has no sharp borders. At one point, you have to stop writing, despite the many untold stories and challenges pinpointed without providing solutions. This is that point. Search engines continue to sink into the things and rhythm of everyday life and to become ever more ambient. They also increasingly merge with other of digital services, where they become just one component in complex algorithmic assemblages increasingly fuelled by artificial intelligence solutions. Future search engines will be even harder to notice for users and researchers alike, but they will be no less significant. This is a challenge for researchers trying to understand the various ways in which people encounter and deal with information and how they envision search technology to work. However, most of all this is a challenge for society and in particular the various professions engaged in mediating, curating, and evaluating information and knowledge claims. The main task ahead is to not shy away from a radical re-examining of notions such as media- and information literacy and to admit to their limitations in the face of the profound societal and infrastructural changes we are faced with and to do this with a sound understanding of the socio-economic and technological arrangements they are embedded in. Not to mystify technology, but to make it visible. Always.

References

Adam, B. (1995). *Timewatch: The social analysis of time*. Cambridge: Polity Press.

Adam, B. (2006). "Time." *Theory, Culture & Society*, 23(2–3), 119–126.

Andrejevic, M. (2013). *Infoglut: How too much information is changing the way we think and know*. New York: Routledge.

Chun, W. H. K. (2008). "The enduring ephemeral, or the future is a memory." *Critical Inquiry*, 35(1), 148–171.

Davies, E. & McKenzie, P. J. (2002). "Time is of the essence: Social theory of time and its implications for LIS research." In *Advancing knowledge: Expanding horizons for information science, 30th annual conference of the Canadian Association for Information Science, Toronto, Ontario, 30 May-1 June, 2002.*

Hartel, J. (2010). "Time as a framework for information science: Insights from the hobby of gourmet cooking." *Information Research*, 15(4), 19. Retrieved from http://InformationR.net/ir/15-4/colis715.html.

Hellsten, I., Leydesdorff, L., & Wouters, P. (2006). "Multiple presents: How search engines rewrite the past." *New Media & Society*, 8(6), 901–924.

Lindsköld, L. (2018). "Google as a political subject: The right to be forgotten debate 2014–2016." *Online Information Review*, 42(6), 768–783.

Moran, C. (2015). "Time as a social practice." *Time & Society*, 24(3), 283–303.

Rosa, H. (2013). *Social acceleration: A new theory of modernity*. New York: Columbia University Press.

Savolainen, R. (2006). "Time as a context of information seeking." *Library & Information Science Research*, 28(1), 110–127.

Thomas, W. I. & Thomas, Swaine D. (1928). *The child in America: Behavior problems and programs*. New York: Knopf.

Willson, M. (2017). "Algorithms (and the) everyday." *Information, Communication & Society*, 20(1), 137–150.

Zavadski, A. & Toepfl, F. (2019). "Querying the Internet as a mnemonic practice: How search engines mediate four types of past events in Russia." *Media, Culture & Society*, 41(1), 21–37.

INDEX