



UvA-DARE (Digital Academic Repository)

The Mysterious User of Research Data

Knitting Together Science and Technology Studies with Information and Computer Science

Gregory, K.; Groth, P.; Scharnhorst, A.; Wyatt, S.

DOI

[10.1007/978-3-031-11108-2_11](https://doi.org/10.1007/978-3-031-11108-2_11)

Publication date

2023

Document Version

Final published version

Published in

Interdisciplinarity in the Scholarly Life Cycle

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Gregory, K., Groth, P., Scharnhorst, A., & Wyatt, S. (2023). The Mysterious User of Research Data: Knitting Together Science and Technology Studies with Information and Computer Science. In K. Bijsterveld, & A. Swinnen (Eds.), *Interdisciplinarity in the Scholarly Life Cycle: Learning by Example in Humanities and Social Science Research* (pp. 191-211). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-11108-2_11

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



The Mysterious User of Research Data: Knitting Together Science and Technology Studies with Information and Computer Science

Kathleen Gregory, Paul Groth, Andrea Scharnhorst,
and Sally Wyatt

K. Gregory (✉)

Visualization and Data Analysis Research Group, Faculty of Computer Science
& Department of Science and Technology Studies, Faculty of Social Sciences,
University of Vienna, Vienna, Austria
e-mail: kathleen.gregory@univie.ac.at

P. Groth

Informatics Institute, Faculteit der Natuurwetenschappen, Wiskunde en
Informatica, University of Amsterdam, Amsterdam, The Netherlands
e-mail: p.t.groth@uva.nl

A. Scharnhorst

Data Archiving and Networked Services, Royal Netherlands Academy
of Arts and Sciences, The Hague, The Netherlands
e-mail: andrea.scharnhorst@dans.knaw.nl

S. Wyatt

Department of Society Studies, Faculty of Arts & Social Sciences,
Maastricht University, Maastricht, The Netherlands
e-mail: sally.wyatt@maastrichtuniversity.nl

© The Author(s) 2023

K. Bijsterveld, A. Swinnen (eds.), *Interdisciplinarity in the Scholarly Life Cycle*,
https://doi.org/10.1007/978-3-031-11108-2_11

Introduction

The open science movement promises to change the production and dissemination of academic knowledge by making the processes and results of research transparent and available, to the benefit of individual researchers, the pocketbooks of funders, and the research enterprise as a whole. Open, accessible, and standardized research data are seen as essential scaffolding for realizing these promises.

Sharing and documenting research data, for example, offer a potential antidote to problems with reproducibility in science by providing a way to validate experimental findings. Reusing research data provides potential economic benefits, by limiting the amount of possibly redundant and nearly always costly data collection. Shared pools of research data offer new possibilities for using data science techniques to tackle society's most wicked problems.

Working to support these visions, data repositories and scientific publishers have increasingly become entangled in mass operations of data documentation and exchange. New tools have been developed to facilitate the discovery of data, and funders and policy makers have implemented policies at national and institutional levels for both open science and data management (European Commission, 2019).

Users are invoked as being central to many of these efforts. Designers of data search tools experiment with sophisticated methods to present the user with the best possible results (e.g., Brickley et al., 2019). Educational tools are designed to help users of repositories and data management tools construct data which are findable, accessible, interoperable, and reusable, or FAIR (Wilkinson et al., 2016). Various metadata schemas, standards, and tools are developed to aid users in discovering and understanding data (e.g., Ohno-Machado et al., 2017).

Despite this stated user focus, the concept of the 'user' or 'users,' similar to that of 'data' and of the practices surrounding data reuse, is conceptualized differently across and within disciplinary domains. In many technical and design-oriented areas of information and computer science, users often remain at arm's length, visible only via ensembles of click behavior, search logs, or data management plans (Van House, 2004).

This acontextual, homogenous view of users contrasts with the heterogeneous, embedded, and socially constructed understanding of use which characterizes work in science and technology studies (STS) (Wyatt, 2003).

Research rooted in these two conflicting views is often undertaken along parallel, yet isolated tracks. When they do intersect, communication between these two perspectives on users is challenging (Tabak, 2014). This chapter reflects upon a project which knit together differing notions of ‘users’ as a way of grounding interdisciplinary research. In addition to producing novel insights about the reuse of research data, this approach also served to bridge the distance between STS researchers and computer scientists, and between designers of data search systems and users themselves.

After explaining the context of the project, we begin from the end, highlighting the results and outcomes which our ‘integrative-synthesis’ approach to interdisciplinarity (Barry et al., 2008) afforded. We then turn to the development of our interdisciplinary approach by exploring the conceptual roots of users within information/computer science and STS and discussing how we wove these ideas together within our research.

We conclude by identifying and reflecting on three points that may be applicable to others conducting interdisciplinary research: (i) a common (yet differently conceptualized) idea, for example, ‘users,’ can serve as an anchor for interdisciplinary work, much in the way of a boundary object; (ii) interdisciplinarity itself is an evolving, contextual construct; and (iii) the broader impacts of interdisciplinary research may change perspectives and practices in ways which are difficult to trace.

Project Re-SEARCH: Contextual Search for Research Data

Project Re-SEARCH was an interdisciplinary project funded by the Dutch Research Council (NWO, grant number 652.001.002) from 2017 to 2021, which brought together industrial and research partners to investigate and develop search solutions for research data. Researchers

from three Dutch universities, a data archive, and an academic publisher pursued three research lines within the project, which were expected to exchange insights and results and eventually resonate with each other. Our STS-infused research line focusing on practices of data discovery and reuse took place alongside research in computer science exploring the development of semantic technologies and relevance ranking algorithms for data search. The academic publisher, Elsevier, provided logistical support for all three lines, with the aim of implementing findings into their prototype search engine for research data, DataSearch.¹

Each research line consisted of senior and junior researchers performing independent research. The search engine design team at Elsevier varied in composition and size over the course of the project, although on average the team consisted of eight individuals, primarily from computer science. The entire team met monthly to discuss how research findings could be implemented into the data search engine, and how data from Elsevier (e.g., search logs and a dataset index) could be used to inform all research lines.

Researchers working on the project came from a variety of disciplinary backgrounds, including computer science and STS, although typically many more computer scientists and system developers than social scientists or humanities scholars were involved. Even within the broad disciplines of information science, computer science, and STS, the project team had many more specific areas of interest and expertise. In our own research line, for example, the junior researcher had a graduate degree in library and information science, and the senior researchers had backgrounds in computer science, STS, philosophy, economics, and physics.

We brought this multiplicity of disciplinary backgrounds to our research questions and aims which sought (i) to explore how researchers across disciplinary domains discover, make sense of, and reuse data which they do not create themselves, and (ii) to inform and intervene into the development of search solutions for research data.

¹ In July 2020, DataSearch was integrated into another Elsevier platform, Mendeley Research Data, available at: <https://data.mendeley.com/research-data/>

Beginning from the End: Main Findings of the Project

Understanding what users do—in this case, how researchers locate data for reuse—was a common interest among all team members. Although both information science/computer science (IS/CS) and STS have long histories of exploring how individuals encounter, understand, and engage with information, so-called information-seeking practices, they have done so from different conceptual and methodological standpoints and have only rarely focused specifically on practices related to data.

In the Re-SEARCH project, we embedded STS perspectives about users, communities, and context into established user-centered models of information-seeking common in IS/CS (further discussed in the next section). We knit these two perspectives into an interdisciplinary theoretical construct which we used to frame a range of quantitative and qualitative methodologies, including an analytical literature review, a large-scale survey, observations, and multiple interview studies. Weaving these two perspectives together led to conclusions about data discovery and reuse, briefly outlined in this section, that both spoke to and challenged traditional notions of users and use, particularly within information and computer science.

One of our principal findings centers on the conceptual development of data communities. In the literature on data discovery and reuse, the term ‘communities’ is often used indiscriminately or to refer to broad disciplinary domains (Borgman, 2012). The results of our research encourage instead a multi-dimensional way of thinking about communities, in which researchers belong to multiple data communities, which are not defined by discipline alone but which rather form around shared data, common data needs, shared methodologies, or common data uses. An example of such a data community can be found in the digital humanities, where researchers from various disciplinary backgrounds come together around a shared (digital) corpus.

Users, of both data and data search systems, are situated within multiple such communities. As they make sense of data for reuse, individuals ‘place’ data within different contexts, for example, contexts of data cre-

ation, disciplinary or social norms, or the data's representativeness of particular phenomena in the world (see Koesten et al., 2021). Our interview study on data-centric sensemaking provides an example of this third type of placing, where study participants worked to place data within the world geographically. Participants questioned if a list of countries in a dataset which we showed them was indeed complete; they also interrogated the granularity of the data, attempting to ascertain if they were representative of an entire country or of only certain areas within a country.

Many of our studies surfaced the importance of data documentation in discovering, placing, and reusing data. Different documentation (e.g., metadata, supporting descriptions, and academic literature) may be needed for different purposes and depend on the 'distances' between users and data in terms of a user's familiarity and expertise. In our interview study focusing on sensemaking, for example, we found that despite their previous knowledge, experienced researchers who are 'close' to the data may need more detailed information than individuals who are 'farther' from the data. This type of detailed information may be best provided using granular, visual representations of patterns rather than the high-level documentation often provided in README files.

We also found that data act as hubs for collaborative activity, as in the case of an early career researcher in the environmental sciences who reported seeking data from other researchers as a way of forming collaborations; data can also provide a means for 'conversations' between data creators and potential reusers. We argue that data discovery systems, repositories, and metadata should be designed to *support* rather than *ignore* the social interactions and collaborative work implicit in discovery and reuse, and we call for innovative solutions, such as interactive forms of data documentation to which both data creators and reusers can contribute.

Finally, our work nuances the idea of 'use,' of both data discovery systems and of data, making visible the multiplicity of actions, resources, and types of data reuse in academic work. We emphasize that both data reuse and the use of discovery systems should be conceptualized as existing on a continuum of uses, rather than as being binary practices of use or non-use. Users of data discovery systems may search for data once or

twice online, for example, or they may look for data routinely. Some researchers only reuse data for teaching; others reuse data for multiple purposes, at multiple timepoints in their work. This suggests that although an individual may at times be a user (of a system or of data), they are also, at times, a non-user.

Knitting Together Differing Notions of Users and Use

The above findings have their theoretical basis in the foundation of our own interdisciplinary approach: the innovative way we knit together the literature on user-centered information-seeking models in information and computer science and the literature on *use*, *communities*, and *context* in STS. We reviewed these literatures to find points of connection and difference and brought them together into a conceptual framework which guided the rest of our research. To provide insight into this framework and its development, we briefly review these two literatures here and examine how we knit them together as a way of grounding our own interdisciplinarity.²

Information and Computer Science Perspective: User-Centered Models

Cognitive, user-centered perspectives to exploring information discovery have held sway in IS/CS since the mid-1980s (Savolainen, 2007) when research into information behavior was seen to have moved from a systems-oriented view to one foregrounding the information seeker's standpoint (Dervin & Nilan, 1986).

These user-centered approaches to studying information-seeking are rooted in the rather indistinct boundaries between computer science and information science, notably in the fields of information behavior, information retrieval, and interactive information retrieval. Although there are

²This section draws heavily on the principal output resulting from our research line (Gregory, 2021).

differences between these fields, they tend to converge on their conceptualization of people, who are defined as ‘users’ of information systems (Jansen & Rieh, 2010).

This view of individuals as users is reflected in numerous conceptual models developed to describe and theorize how people seek information. Such models, many of which date from the final decades of the twentieth century, usually consist of diagrams describing relationships among concepts (Case & Given, 2016), for example, individuals, systems, information, and actions. The majority focus on *search* behaviors (Järvelin & Wilson, 2003), examining how a user interacts with a search system to satisfy an information need.

Information-seeking models have varying levels of specification and serve different research purposes, for example, interpreting observations (Järvelin & Wilson, 2003) or investigating search practices of particular groups (e.g., Ellis & Haugan, 1997). The information journey model (Blandford & Atfield, 2010), which particularly informed our work, synthesizes many key aspects of earlier models. In this model, an information seeker moves through four stages, which are not necessarily required or sequential: *recognizing* a need for information; *acquiring* information, either through active searching, serendipitous discovery, or being told about it; *interpreting and evaluating* information; and *using* the interpreted information.

Across user-centered models, including the information journey model, information discovery is defined through interactions between users and systems but also between contexts and users. Context is a complex and variously defined concept in information research. Positivist views portray context as a backdrop for activities, as an itemized yet inexhaustible list of elements, whereas more relational views see context as being an enacted and local ‘carrier of meaning’ (Dervin, 1997.) With some exceptions (e.g., Saracevic, 1996; Ingwersen, 1996), established information-seeking models tend toward positivist conceptions, where context is composed of nameable cognitive and affective factors (e.g., goals, tasks, prior knowledge, or feelings such as optimism or uncertainty) which influence a user’s information activities (Courtright, 2007). Although many user-centered models start from the view that context

shapes search behaviors, they do not address how search behaviors affect broader contexts.

Nancy Van House (2004) summarizes many of the limitations of how users are portrayed in models of information-seeking. Perhaps obviously, user-centered models focus on the viewpoint of one individual acting in a particular role: the person looking for information. This person is defined in terms of interactions with an information system. Other roles that a person enacts, which may influence information practices (IP), are not addressed. Individuals who do not interact with systems are not typically represented. User-centered models do not meaningfully draw out other actors involved in information discovery processes; nor do they account for shared or distributed actions (Talja & Hansen, 2006).

STS Perspective: Use, Communities, and Context

The division we make between IS/CS and STS might seem a bit artificial. STS-inspired approaches have been taken up within sub-fields of information science, most notably in the area of ‘information practices’ (IP) research, which examines how information-seeking activities are influenced and shaped by both social and cultural factors (Tuominen et al., 2005), and where information-seeking and use are seen as constructing activities. This area of research reflects the STS tenet that society and technology are not separate entities but are instead co-constituents of a seamless web of dynamic social, material, political, and economic elements (van House, 2004).

In IP research, the focus is not on users per se but rather on the socio-technical infrastructures, practices, and contextual factors surrounding and shaping information-seeking (Savolainen, 2007). The term ‘users’ is also avoided or invoked with care. For example, social informatics proposes the term ‘social actors’ to recognize that individuals using technologies are not primarily defined by that use but rather enact multiple roles and inhabit multiple contexts where technologies are present (Lamb & Kling, 2003).

The treatment of users in IP research is mirrored in STS, which has a history of examining the practices of scientists, engineers, and

technologists, rather than those of 'users.' When users of a technology are studied, they are viewed as heterogeneous, embedded in dynamic, locally enacted contexts, and are investigated symmetrically. Studies investigate not only how technology shapes user practices but also how users shape the development of technologies (Oudshoorn & Pinch, 2003).

Focusing only on those who engage with a technology ignores the people who do not and thus reinforces the idea that technology use is the norm (Wyatt, 2003). As with users, non-users are not a homogenous group. As Sally Wyatt points out, people may not have access to a technology, they may actively choose not to engage with it, or they may have tried it once or twice and decided that it was not for them. Users should therefore be "conceptualized along a continuum, with degrees and forms of participation that can change" (Wyatt, 2003, p. 77).

Much work in STS also argues that the community or collective, rather than the individual, is the entity that 'knows' (van House, 2004). Knowledge production is situated within these collectives, which have unique norms, practices, and tools. The actions of an individual are representative of these collective ways of knowing and producing knowledge (Knorr Cetina, 1999); the study of individual actors/'users' is therefore inextricably linked with the study of communities. The embedding of individuals and knowledge in communities is also present in information practices research, which emphasizes that individuals act as members of communities and social groups, often in diverse roles, rather than as isolated actors, and that information-seeking is a social practice (Talja & Hansen, 2006).

Work by Karin Knorr Cetina on epistemic cultures (1999) is representative of core ideas about how STS conceptualizes communities. Communities are seen as dynamic groups that are forged through common practices, epistemic norms, and shared objects. Disciplinary domains alone do not define communities. Individuals can belong to multiple communities, and individual actions are shaped by communities themselves. The key to understanding these communities lies in studying practices to reveal actual, rather than imagined, actions and relationships.

Interdisciplinarity in Practice

In Project Re-SEARCH, we brought the STS perspective described above into user-centered models of information-seeking. A slightly modified version of the information journey model (Blandford & Attfield, 2010), adapted to reflect our research questions, structured our empirical studies. This modified version highlights *users and their needs* and their practices of *discovering*, *evaluating/sensemaking*, and *(re)using* information, or in the case of our research questions, data. We also drew on tenets of other models (e.g., Saracevic, 1996; Ingwersen, 1996) to understand data discovery as a dynamic process, shaped by a user's purpose or task and involving multiple strategies over time.

Rather than situating data discovery and reuse as isolated practices, we drew on the STS perspectives reviewed above to emphasize the embedding of practice within dynamic networks of people, technologies, materials, and policies. We viewed users not as atomized individuals but rather as social actors enfolded in epistemic communities who engage with technologies in various ways. We worked from the idea that examining objects of research from different perspectives could serve to uncover actual practices of use, to reveal the diversity and multiplicity of practice, and to explore relationships between sociotechnical elements.

To operationalize this perspective, we developed a series of guiding questions (Fig. 1, center). Identifying the various communities, types of data, and technologies used was an important step in understanding data discovery practices. This involved asking 'which' questions, that is, 'which users,' 'which communities,' or 'which data'; it also involved paying attention to which entities were not being taken up or which communities were not engaging in a practice, for example, by identifying absent communities in the current literature or by paying attention to non-response in our survey analysis (Q1). Studying practices of data discovery and reuse required observing what people were doing, as well as questioning how these practices were changing or stabilizing in relation to technologies, materialities, and norms (Q2). Understanding the motivations underlying and the consequences of practice helped to provide context and to trace relationships between particular practices and other elements (Q3, Q4).

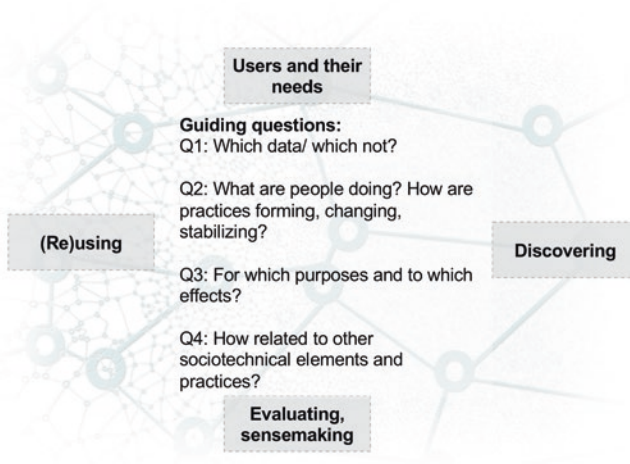


Fig. 1 Representation of integrated theoretical approach. Boxes represent modified sections of the information journey model from the IS/CS perspective. Guiding questions were developed from key points about use, communities, and context from an STS perspective

Figure 1 visualizes how our conceptual approach was informed by both models and concepts from IS/CS and STS. We further used this interdisciplinary conceptual framework to structure and analyze our empirical work. For example, in our initial study, we conducted an analytical literature review structured along the lines of common features of information-seeking models: user needs, search strategies, and evaluation criteria. At the same time, we analyzed the literature through the lens of a multi-dimensional approach to understanding communities. Rather than focusing on individual data users, we looked for commonalities in practices which might lead to new, emerging groupings of those ‘users,’ for example, around different types of data.

We conducted two further studies: semi-structured interviews with data seekers and a large-scale global survey. We used the anatomy of the modified information journey model to organize both the interview protocol and the survey questionnaire. In follow-up questions during the interviews and in our analysis of both the quantitative and qualitative

survey data, we raised questions about what was present in researchers' data practices. We also sought to identify communities and practices which were not as visible and to tease out relationships between data discovery and other research practices. For example, we asked research participants to explain the similarities, differences, and overlaps between discovering and understanding academic literature versus data, and we encouraged them to explore and untangle the role of professional and personal networks in data discovery and reuse.

Making Interdisciplinarity Visible

Our interdisciplinary approach is made tangible in the craftwork which we have included at the beginning of this chapter. This data visualization, a 'Knitted Web of Science,' represents the references cited in the principal research output of our research (Gregory, 2021). Wyatt classified all the references, according to her knowledge of the work, and/or by the journal. She then calculated the proportion of the total references for each of the eight disciplinary domains.

These disciplines are knitted in the usual order, starting with the largest, green for information science. Other disciplinary domains in descending order are other natural and life sciences, FAIR and open data, computer science, libraries/archives, methods, STS, and other social sciences and humanities. Our personal areas of expertise are also represented with buttons. The three larger buttons represent the main expertise of Groth, Scharnhorst, and Wyatt. The smaller, pearl buttons spread across the visualization represent the multiple venues where Gregory's work has been published and her command of different disciplinary repertoires.

Lessons Learned About Interdisciplinary Research

The 'knitting together' of perspectives and methods represented in this craftwork helped us to expose new interdependencies and reach conclusions that would have been difficult to arrive at if we had relied on only

one perspective. It also allowed us to apply our findings to develop recommendations for systems design and data documentation and to communicate those to individuals from various (disciplinary) backgrounds via the language of ‘users’ and ‘uses,’ which resonated with a variety of audiences. A perhaps unexpected outcome from this research was the opportunity it provided to reflect on our own understanding of interdisciplinarity and to identify points for others to consider when conducting interdisciplinary research.

A Shared (Yet Differently Conceptualized) Term Is Often an Anchor

IS/CS and STS have approached studying use and users differently, both conceptually and methodologically. It was these differing notions about a common concept that served to ground our interdisciplinary research.

As a concept, ‘users’ are widely employed within modern software development, where user experience design and testing play a fundamental role (Kashfi et al., 2017). This provided us with a connection point not only with other team members but also with the software developers on DataSearch and with other developers, for example, at the data archive where two of us (Gregory and Scharnhorst) were employed.

The idea of ‘users’ acting as an anchor point for our work has parallels with Peter Galison’s metaphor of the ‘trading zone,’ where individuals from potentially incompatible viewpoints develop a shared language as a way of bridging differences (Galison, 1997). In our research, we did not match different terminologies or define a shared, working vocabulary, as sometimes happens. Instead, we worked to make explicit the different connotations of a single term. This can be important for interdisciplinary collaboration: if these differences are not made explicit, confusion can arise when people think they are talking about the same concept.

The concepts of ‘users’ and ‘use’ acted more as boundary objects in our work (Star & Griesemer, 1989), traveling between different disciplinary and professional communities while being malleable enough to be adapted to fit our research questions and project aims. We then communicated our findings through the frame the ‘user’ boundary object

provided, relying on this concept to reach and resonate with other research and practitioner communities.

Interdisciplinarity Exists in Different Contexts and Evolves over Time

Not all forms of interdisciplinary research are the same. Interdisciplinarity is enacted through different relationships between contributing disciplines and by various levels of engagement. Andrew Barry and colleagues propose three different modes of interdisciplinarity (Barry et al., 2008). The *integrative-synthesis* mode is characterized by a roughly symmetrical integration of methods and concepts from the involved disciplines. In *service-subordination* mode, one discipline is seen as existing in service to others, contributing without significantly changing the rules of other disciplines (Wyatt, 2021). The *agonistic-antagonistic* mode captures one discipline explicitly aiming to change another. As we saw in project Re-SEARCH, these modes of interdisciplinarity can co-exist and be perceived differently within the same project.

We viewed our research as integrative-synthesis work, bringing together different perspectives from IS/CS and STS in order to enrich the entire project. Our research partners may have had a different view of our role, particularly at the beginning of the project. It could be argued that Elsevier initially saw our STS research line as existing in a service-subordination relationship, where we were expected to adopt the ‘correct objective’ of the project as a whole, namely to focus on how potential and future users/researchers interact with the DataSearch search engine. At the same time, without Elsevier’s engagement, we would not have been able to conduct our own research, particularly the survey, at the same scale or with the same populations.

As the project proceeded, our work challenged the service-subordination model. We did not limit our research to interactions with DataSearch but rather expanded the scope of the problem to focus on data discovery and data reuse practices more broadly, in addition to uses of search technologies. Our STS focus shifted our research line from being object-oriented, focusing on DataSearch, to being practice-oriented, where ‘users’

provided an entry point to the wider universe of technologies and practices implicated in data discovery and reuse.

Reflecting this, our partners' views of our work evolved as the project progressed. As we deepened our integration of IS/CS and STS perspectives in our own research and as we began to publish our findings and accrued more data, other partners came to see us on a more equal footing. This shift was perhaps accelerated by the fact that we published our work in journals in information science and data science, some of which our project partners had published in themselves. We also collected and analyzed a substantial amount of data in the survey study. These data and our quantitative analysis were perhaps more closely aligned with other partners' own conceptions of what constitutes high quality research. These similarities may have shifted perceptions about our contributions and the role of our research.

Tracing the Effects of Interdisciplinarity Can Be Challenging

Discussions about tracing the 'impact' of academic work often turn to measures which are easily visible, such as citations. In the short time since we completed our research, we already see signs of these traces. Our studies have been cited in multidisciplinary and discipline-specific journals and in relation to different topics, for example, systems development and data stewardship. Citation practices in different sources, disciplines, and communities vary, which can make it difficult to find and place such citations in the correct contexts.

Another way of viewing the impact of our project could be in the implementation of our recommendations in search solutions for research data, particularly in DataSearch. Changes have been made which align with our recommendations, such as indexing a wider diversity of data repositories and providing links to literature databases. It is difficult to say, however, whether these developments were always planned or whether our research directly contributed to the system's development.

It may take time for the impact of our work to become traceable via such mechanisms. We argue that our work has produced more subtle

shifts in practices and perspectives, both in our own team and among others, which are not as visible as impacts documented through citations or new system features.

Signs of these types of impact can be seen in the various invitations for talks and workshop participation which we have received and, also, in discussions around the development of other data search systems, such as Google's Dataset Search. For example, Paul Groth gave a talk on data reuse at the Chan Zuckerberg Initiative; Kathleen Gregory was invited to participate in a Dagstuhl computer science workshop on FAIR data infrastructures. After publishing our survey results, we had conversations with Google Dataset Search about a possible collaboration. Although we did not enter a formal collaboration, a similar discourse to ours can be seen in recent descriptions about Google Dataset Search (e.g., Brickley et al., 2019).

Various team members of Project Re-SEARCH have also reported that their way of viewing data reuse has been altered by our work. Within data archives, especially at the Data Archiving and Networked Services (DANS) where Gregory and Scharnhorst were/are employed, experimenting with different ways of understanding 'users,' particularly as members of data communities, has provided stimulation for moving beyond the idea of organizing archival services only along disciplinary groups.

These less formal signals are indications that our form of interdisciplinarity has helped to bridge the distance between STS researchers and computer scientists, and between designers of data search systems and users. Our approach also has the potential to shape views about 'users' within STS. For example, the models of information-seeking from information science which we drew on, could help to attune STS to different types and temporalities of use. Furthermore, the 'Knitted Web of Science' which we used to introduce and illustrate this chapter (Fig. 2) makes the extent of our interdisciplinary collaboration visible. Knitting one's references may not be a route all researchers could or should adopt, but some awareness of and attention to one's literature and citation practices could help us all to expand our horizons.



Fig. 2 The 'Knitted Web of Science' © Gregory et al.

References

- Barry, A., Born, G., & Weszkalnys, G. (2008). Logics of interdisciplinarity. *Economy and Society*, 37(1), 20–49. <https://doi.org/10.1080/03085140701760841>
- Blandford, A., & Attfield, S. (2010). *Interacting with information*. Morgan & Claypool Publishers. <https://doi.org/10.2200/S00227ED1V01Y200911HCI006>
- Brickley, D., Burgess, M., & Noy, N. (2019). Google Dataset Search: Building a search engine for datasets in an open Web ecosystem. *The World Wide Web Conference-WWW '19*, 1365–1375. <https://doi.org/10.1145/3308558.3313685>
- Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078. <https://doi.org/10.1002/asi.22634>
- Case, D. O., & Given, L. M. (2016). *Looking for information: A survey of research on information seeking, needs, and behavior* (4th ed.). Emerald Group Publishing.
- Courtright, C. (2007). Context in information behavior research. *Annual Review of Information Science and Technology*, 41(1), 273–306. <https://doi.org/10.1002/aris.2007.1440410113>
- Dervin, B. (1997). Given a context by any other name: Methodological tools for taming the unruly beast. In P. Hakkari, R. Savolainen, & B. Dervin (Eds.), *Information seeking in context* (pp. 13–38). Taylor Graham.

- Dervin, B., & Nilan, M. (1986). Information needs and uses. *Annual Review of Information Science and Technology*, 21, 3–33.
- Ellis, D., & Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, 53(4), 384–403. <https://doi.org/10.1108/EUM0000000007204>
- European Commission. (2019). *Facts and figures for open research data*. https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/facts-and-figures-open-research-data_en
- Galison, P. (1997). *Image and logic: A material culture of microphysics*. University of Chicago Press.
- Gregory, K. (2021). *Findable and reusable?: Data discovery practices in research*. [Doctoral dissertation, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20210302kg>
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), 3–50.
- 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*. <https://doi.org/10.1002/asi.21358>
- Järvelin, K., & Wilson, T. D. (2003). On conceptual models for information seeking and retrieval research. *Information Research*, 9(1), 1–23.
- Kashfi, P., Nilsson, A., & Feldt, R. (2017). Integrating User eXperience practices into software development processes: Implications of the UX characteristics. *PeerJ Computer Science*, 3, e130. <https://doi.org/10.7717/peerj-cs.130>
- Koesten, L., Gregory, K., Groth, P., & Simperl, E. (2021). Talking datasets – Understanding data sensemaking behaviors. *International Journal of Human-Computer Studies*, 146, 102562. <https://doi.org/10.1016/j.ijhcs.2020.102562>
- Knorr Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.
- Lamb, R., & Kling, R. (2003). Reconceptualizing users as social actors in information systems research. *MIS Quarterly*, 27(2), 197–235.
- Ohno-Machado, L., Sansone, S., et al. (2017). Finding useful data across multiple biomedical data repositories using DataMed. *Nature Genetics*, 49(6), 4–7.
- Oudshoorn, N., & Pinch, T. (2003). Introduction: How users and non-users matter. In N. Oudshoorn & T. Pinch (Eds.), *How users matter. The co-construction of users and technology* (pp. 1–25). The MIT Press.

- Saracevic, T. (1996). Modeling interaction in information retrieval (IR): A review and proposal. *Proceedings of the 59th Annual Meeting of the American Society for Information Science*, 33, 3–9.
- Savolainen, R. (2007). Information behavior and information practice: Reviewing the “umbrella concepts” of information-seeking studies. *The Library Quarterly*, 77(2), 109–132. <https://doi.org/10.1086/517840>
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, ‘translations’ and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, 19(3), 387–420. <https://doi.org/10.1177/030631289019003001>
- Tabak, E. (2014). Jumping between context and users: A difficulty in tracing information practices. *Journal of the Association for Information Science and Technology*, 10.
- Talja, S., & Hansen, P. (2006). Information sharing. In A. Spink & C. Cole (Eds.), *New directions in human information behavior* (pp. 113–134). Springer Netherlands. https://doi.org/10.1007/1-4020-3670-1_7
- Tuominen, K., Talja, S., & Savolainen, R. (2005). The social constructionist viewpoint on information practices. In K. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of information behavior* (pp. 328–333). Information Today.
- Van House, N. (2004). Science and technology studies and information studies. *Annual Review of Information Science and Technology*, 38, 3–86.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wyatt, S. (2003). Non-users also matter: The construction of users and non-users of the Internet. In N. Oudshoorn & T. Pinch (Eds.), *How users matter: The co-construction of users and technology* (pp. 67–79). The MIT Press.
- Wyatt, S. (2021). Interdisciplinarity: Models and values for digital humanism. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on digital humanism* (pp. 329–333). Springer.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

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

