

Research Protocol

Overview

This study investigates the social, technical, administrative and epistemic factors that mediate data harmonization initiatives in epidemiological research. It takes to heart the notions that data are media that facilitate communication across different research contexts, that data are created with specific intent, and that data are bounded by the social, practical and material circumstances of their creation. In light of these facts, I approach data harmonization as means of reconciling the varied circumstances of datasets' creation — both among themselves, and in relation to contexts of reuse.

I therefore frame data-sharing as initiatives to foster a series of collaborative ties beyond a project's original intended scope. Data harmonization is a means of drawing the observations that constitute data into a formal schema, whose structure is driven by specific objectives and more general underlying suppositions and values. Moreover, schemas are arrived at through discussion, compromise and consensus-building, with an eye toward practical outcomes afforded by the data model.

While data harmonization certainly benefit from discursive interactions, these are generally under-valued as important factors contributing to data-sharing initiatives' success. My goal is to survey what factors are being prioritized at various data harmonization initiatives, the rationales behind these decisions, and the relative efficacy of these variable approaches.

To be clear, my intent is not to pit various approaches against each other, but rather to ascertain what actions entail, the circumstances in which each is adopted, the value that they bring, and the trade-offs involved. In other words, I want to reveal the diverse ways in which data-sharing occurs, and how different approaches impact the outcomes in different ways.

The Maelstrom Project, which is a leading firm supporting data-sharing initiatives in epidemiology research, presents a great opportunity to explore how social and material factors are being accounted for in data harmonization efforts. Maelstrom operates by partering with research projects through initial consultations, which may then evolve into more comprehensive data harmonization work. This is contingent on the value proposition that Maelstrom and partner projects ascertain will derive from harmonization, and evaluation of the feasibility of achieving these outcomes. Already, this approach differentiates itself from “raw”¹ open data-sharing in that it is directed by specific objectives, recognizes limitations of practical circumstances surrounding data's creation and the data harmonization efforts, and maintains the option to not proceed if it is deemed prudent to do so. Maelstrom's partners represent a pool of potential cases that already grapple with issues concerning mediation of situated experiences in data-sharing, and which may be receptive to investigation of their research practices.

Relevance

This project is directly relevant to ongoing developments in the open data landscape. As the open data movement begins to mature, cracks are beginning to reveal themselves in the infrastructures we have built thus far. Critical inspection is therefore necessary to help improve these systems and ensure that they may continue to support research activities.

Moreover, as a case study on community-oriented data-sharing initiatives, this project is well equipped to draw attention to the support structures (or lack thereof) for these efforts. Specifically, the project will contribute to a better understanding of what resources are necessary to improve data-sharing at large and small scales.

¹I tentatively use the term “raw” data-sharing to mean acts of uploading and downloading spreadsheets among strangers via the web, which I tend to characterize as transactional (rather than commensal), as oriented toward compliance with the emerging bureaucratization of open science, and as relatively asocial in nature. See my [blog post](#) where I rant about this in greater depth.

This is especially relevant in the Canadian context, where open science policy has been undergoing major revisions for several years now, and which inspires little confidence in researchers concerning expected outcomes. Researchers have therefore taken it upon themselves to develop data-sharing initiatives on their own terms. This entrepreneurialism has been a boon for community-driven data-sharing, but is also plagued by difficulties, which this project will be the first of its kind to explore.

Research Questions

Still in progress...

Methodology

Approach

This study is informed by a set of theoretical and methodological frameworks formed within a more interdisciplinary “science studies” tradition, which contribute to a more sociological outlook on science as cultural practice (cf. Pickering 1992). In practical terms, I will document the social and collaborative experiences involved in various research practices, which ultimately bind the many ways in which scientists do science.

I will specifically focus on how people contribute to and extract from information commons, which comprise both formal documents and mutually-held and information-laden situated experiences. This involves examining the ways in which participation in disciplinary or even more specialized communities of practice fosters mutual understanding about the potential and limitations pertaining to other people’s data; and how this communally-held knowledge is accessed and re-produced. This approach aligns with the situated cognition methodological framework for examining the improvised, contingent and embodied experiences of human activity, including science (cf. Suchman 2007; Haraway 1988).

The situated cognition framework prioritizes subjects’ outlooks, which are contextualized by their prior experiences, and enables scholars to trace how people make sense of their environments and work with the physical and conceptual tools available to them to resolve immediate challenges. Situated cognition therefore lends itself to investigating rather fluid, open-ended and affect-oriented actions, and is geared towards understanding how actors draw from their prior experiences to navigate unique situations.

Situated cognition is especially salient in explorations of how people who are learning new skills learn how to work in new and possibly unfamiliar ways, and in this sense is closely related to Lave and Wenger’s (1991) theory of situated learning (or ‘communities of practice’ approach), which focuses on how individuals acquire professional skills in relation to their social environments. In such situations, situated cognition enables observers to examine how people align their perspectives as work progresses, and to understand better how people’s general outlooks may have changed under the guidance of more experienced mentors. In other words, situated cognition enables researchers of scientific practices to account for discursive aspects of work, including perceived relationships, distinctions or intersections between practices that professional or research communities deem acceptable and unacceptable, and the cultural or community-driven aspects of decisions that underlie particular actions.

In taking on this theoretical framework, I frame epidemiology as a collective endeavour to derive a coherent understanding of population-level health trends, which involves the use of already established knowledge in the validation of newly formed ideas, and which relies on systems designed to carry information obtained with different chains of inference. These systems have both technical and social elements. The technical elements are the means through which information becomes encoded onto information objects so that they may form the basis for further inference. The social elements constitute a series of norms or expectations that facilitate the delegation of roles and responsibilities among agents who contribute their time, effort and accumulated knowledge to communal goals.

As such, in constructing the arguments of this study and in carrying out the fieldwork that grounds it, I will rely upon both realist and constructivist viewpoints. In one sense, I rely on documenting how people actually act, including the longer-term and collaborative implications that their actions may have on other work occurring throughout the continuum of practice. To accomplish this, I identify research activities from

the perspective of an outside observer. I also ascribe meanings to things (such as physical or conceptual tools, or objects that captivate subjects' interests) in ways that conform to my own perspective as an investigator of scientific research practices.

On the other hand, a constructionist perspective enables me to consider how individual agents make components of information systems suit their needs to facilitate communication or interoperability among actors who hold different situated perspectives. By listening to participants' views about the systems with which they engage, including explanations as to why they act in the ways that they do, I am able to trace the assumptions and taken-for-granted behaviours that frame their perspectives. Moreover, these insights are useful for developing a better understanding of how participants identify with particular disciplinary communities and their perception of their roles within broader collective efforts.

Ultimately, this study is about the social order of scientific research, i.e. the frameworks, mindsets or sets of values that humans adopt to carry out their work in specific ways. Human beings rely upon physical and conceptual apparatus to do this work but, in order to understand how they do science *in ways that conform to the epistemic mandates of the scientific enterprise*, it is necessary to prioritize attention to human intention, drivers and pressures. I am emphasizing the agency of human drivers – as opposed to tools and procedures² – since humans are the ones who (a) identify problems that need to be resolved; (b) imagine, project or predict potential outcomes of various kinds of actions that they may select to resolve the challenges; and (c) learn from prior experiences and change their behaviours accordingly. By highlighting how pragmatic actions are conducted in relation to broader social and discursive trends and tendencies, I consider scholarly practices in terms of potential, certainty and desire from the perspectives of practitioners themselves.

Data

This paper will draw from interviews with epidemiological researchers and those who support epidemiological data integration. Interviews will be driven by an emphasis on documenting the process of reconciling different stakeholders' interests as they converge in the formation of a common data stream. Specifically, interviews will focus on motivations for their data harmonization work, the challenges they experience, how they envision success and failure, their perceptions of their own roles and the roles of other stakeholders, the values that inform their decisions, how the technological apparatus they set up enables them to realize their goals and values, and things they consider to be missing from data-sharing initiatives they have experienced thus far.

I will hold around 12-15 interviews and analyze participants' elicitations using qualitative data analysis methods. Interviewees will comprise members of several projects that partner with the Maelstrom research group, a service that supports data harmonization efforts in the field of epidemiology. My selection of cases reflects a need to account for a wide range of decisions and values accounted for in data harmonization efforts.

Data collection will be facilitated by the fact that many of the projects are based in Canada, and specifically in Montreal – where the principal researcher currently resides. This may introduce some bias toward prioritizing local connections (as opposed to Maelstrom's work with European- and American-led projects) and increased visibility toward collaborative benefits that accompany local cooperation. However, this seems to be intrinsic to Maelstrom's roots, and may not really be an aberration.

I will record all interviews using a SONY ICD-UX560 audio recorder to capture audio in the lossless 16 bit 44.1 kHz Linear PCM wav format. I may apply additional audio filters to enhance audio playback while transcribing selected sections of interview records, depending on the degree of noise interference present during the interviews. I will also record interview sessions using a GoPro Hero 4 Silver action camera, pending interviewees' informed consent. Based on prior interviews with scientists about their research experiences, I found that interviewees like to show me, rather than merely tell me, about what they are working on and the means through which they engage with information systems. The action cameras may be leveraged to record spontaneous video records of these demonstrations and provide me with an additional

²Human and non-human agents are considered on equal footing under the Actor-Network Theory (ANT) framework, which has gained some traction since its origins in the late 1980s, but which may not be suitable for this approach. See the appendix for further commentary on ANT and its alternatives.

rich data source for further analysis. Moreover, the action cameras provide an additional backup audio recording in case of data loss on the primary recording device.

I will record handwritten notes in bound notebooks. They will largely comprise descriptive documentation of activities and interactions when recording devices are not operating, as well as preliminary interpretations of observed behaviours and notes on things I plan to follow up on at a later time. I will scan these notebooks and transcribe the text so that these additional data sources could be included in the analytical corpus.

I will transcribing the interviews and edit the transcripts so that they conform with transcript notation standards (e.g. GAT-2, see Selting, Auer, and Barth-Weingarten 2011) and so they are optimally formatted for my analytical intent. I will use automated speech recognition and natural language processing to create preliminary transcripts, which I will then manual edit. I may leverage support from a research assistant for these tasks.

I will collect and handle all data in full compliance with an ethics protocol to be approved by McGill University’s research ethics board. I will present an overview of the research to all participants and obtain verbal consent prior to each interview before proceeding. I will opt to not render confidential the identities of study participants, but I will provide them with the option if they desire. This will include the ability to retroactively select this option up to one week after the interview has been completed. I will prompt interviewees about this option immediately after the termination of the interview and via email the day before the end of the one week period following the interview.

Something about data management plan [in progress]...

Methods

I will qualitatively analyze transcribed interviews to highlight the collaborative nature of work involved in harmonization of epidemiological data, which participants may struggle to identify or articulate. My analysis will involve an abductive qualitative data analysis methodology to articulate theories based on empirical evidence. This approach draws from what Charmaz (2014, 14–15) calls the “constellation of methods” associated with grounded theory that are helpful for making sense of qualitative data, namely coding and memoing.

Coding, which involves defining what data are about in terms that are relevant to the theoretical frameworks that inform my research, entails rendering instances within a text as interpreted abstractions called codes (Charmaz 2014, 43). Codes can exist at various levels of abstraction. For instance, I may apply descriptive codes to characterize literal facets of an instance within a text, and theoretical codes to represent more interpretive concepts that correspond with aspects of particular theoretical frameworks. I may create many of my codes on the fly as “open codes” when prompted by encounters with demonstrative instances in the text. As I create new codes, I will situate them within a code system that provides me with a rough taxonomic structure to help organize my work and to enable me to more effectively query the data. Coding in this manner involves synthesis of concepts that speak to my understanding of the phenomena of interest, while forcing me to remain receptive to limits imposed by what is actually contained in the text. In other words, coding involves applying a precise language to segments of transcribed interviews that serve to bridge the gap between what participants said and the theoretical frameworks I applied to explore them as epistemic activities, interfaces and values (cf. Charmaz 2014; Saldaña 2011, 95–98).

Memoing entails more open-ended exploration and reflection upon latent ideas in order to crystallize them into new avenues to pursue (Charmaz 2014, 72). Constructing memos is a relatively flexible way of engaging with data and serves as fertile ground for honing new ideas. Memoing is especially crucial while articulating sensitizing concepts, which Charmaz (2003, 259) refers to as the “points of departure from which to study the data”. Memoing allows me to take initial notions that lack specification of well-defined attributes, and gradually refine them into more cohesive, definitive concepts (Blumer 1954, 7; Bowen 2006). Exploring the main features, relationships or arrangements that underlie a superficial view of a sensitizing concept through memoing helps me to identify what kinds of things I need to locate in the data in order to gain a full understanding of the phenomena of interest. Memoing is also very important in the process of drawing out more coherent meaning from coded data (cf. Charmaz 2014, 181, 290–93). By creating memos pertaining to

the intersections of various codes and drawing comparisons across similarly coded instances, I am able to form more robust and generalizable arguments about the phenomena of interest and relate them to alternative perspectives expressed by others.

Throughout my analysis, I will follow the approach that Nicolini (2009) and Maryl et al. (2020, para. 30) advocate, who suggest “zooming in to a granular study of particular research activities and operations and zooming out to considering broader sociotechnical and cultural factors”. This involves “magnifying or blowing up the details of practice, switching theoretical lenses, and selective re-positioning so that certain aspects are fore-grounded and others are temporarily sent to the background” (Nicolini 2009, 1412). This approach is useful for me because research projects all start from different positions but share common practices and tendencies that vary according to those contextual circumstances. I am therefore able to tactfully switch between those lenses to understand the interplay between circumstances and practical implementations, which vary across cases, which have their own histories, memberships, sets of tools, methods, and social or political circumstances. ships, sets of tools, methods, and social or political circumstances.

I will perform all of this work using MaxQDA, a proprietary qualitative data analysis software suite that stores all of these connections within a centralized database (VERBI Software 2021).³ This allows me to retrieve segments of text from across various documents that have been assigned the same sets of codes, and perform more complex queries that search along different parameters of overlap, intersection, and exclusion. I will then able to query the integrated dataset to produce elaborated accounts of specific kinds of activities, decisions, values and sentiments.

Limitations

It should be noted that case study research limits my ability to derive generalized findings across the whole field of epidemiology. As such, my intent is instead to articulate some significant aspects of data harmonization as they are represented at the cases accessible through this study. In other words, I aim to make certain under-appreciated social and collaborative commitments that underlie data-sharing initiatives more visible and to draw greater attention to certain sensibilities, attitudes, and apprehensions that are relevant to contemporary discourse on the nature of epidemiological data and ongoing development of information infrastructures designed to support data integration and re-use.

Still in progress...

- Potential difficulties entering the community

Expected Outcomes and Impact

Still in progress...

- Support efforts to enhance data reuse potential of the CITF Databank
- Devise and publish practical protocols for developing community-centred open data-sharing initiatives (maybe a “10 simple rules” style article?)
- Submit comments to CIHR with regards to their ongoing revisions to federal open access policy
- Publishing in STS/meta-science and epidemiology journals (specific venues TBD)

Appendix

Comments on Actor-Network Theory

Previous work examining the apparatus of science relies heavily on Actor-Network Theory (ANT) or similar approaches that highlight the agency and impacts of tools on research practice. These draw attention to how non-human objects not only frame how human beings inhabit the world, but “push back” upon human actions with significant effects (Latour 1992). In the context of early studies of scientific research, ANT

³I am most familiar with MaxQDA but I may explore other software. NVivo is another popular proprietary software suite with overlapping functionality. QualCoder, OpenQDA and QCoder are open source but they are under ongoing development, and may not be as feature-rich as proprietary systems once analysis commences.

was used to understand the physical and communicative mechanisms – made up of non-human agents and information objects – upon which scientists rely to capture, document and ascribe meaning to particular facets of the world (cf. Latour and Woolgar 1986). ANT posits that scientists can only identify, characterize and understand objects of interest by co-creating their conceptions of reality alongside non-human agents.

For instance, (Latour and Woolgar 1986) initially applied ANT to investigate how knowledge is produced by molecular biologists in a laboratory setting. In this work, they identified the processes of transcription and translation as mechanisms through which material engagements between non-human objects (i.e. objects of nature) are rendered in ways that may be understood by humans. For instance, in a biomolecular lab, a bioassay determines and represents the existence and qualities of discrete proteins that have been extracted from cell samples through reactions with specially selected reagents, marked with dyes selected for their non-reactive qualities, and pulled through a filtration medium that separates proteins according to their mass, molecular structure, and electromagnetic properties. A large number of material entities are thus selected and mobilized due to the researcher's understanding that such an assemblage might contribute to the representation of a targeted set of phenomena. More specifically, certain reagents are added to cell samples that bind with proteins that constitute cell walls, allowing for inner proteins to be extracted. Other reagents that only bind to particular targeted proteins and that have been engineered to have certain identifiable physical and electromagnetic properties impart these qualities onto the now merged molecule composed of both targeted protein and reagent. Dyes that complement the design of the reagent and of the overall bioassay protocol are bound to these molecules, rendering them visible to the human eye. Molecules are then pulled through a conductive gel whose own material characteristics and density allow for only electrically charged molecules to be pulled through, given the placement of a polar-opposite charge at the far end of the gel. Once movement of the proteins have settled, (caused, in part, by the gradual solidification of the gel that inhibits further movement) the entire gel is photographed and the relative positions of dyed molecules are marked to denote their discrete existence and other relevant qualities rendered through association with their respective reagents. Knowledge was thus obtained by humans through the selective assembly of various material actors, whose interactions are already understood and deemed reliable, in order to render an account of reality based upon their projected experiences.

Such a comprehensive view of the system of interacting agents, however, betrays the pragmatic and situated nature of the process. Indeed, Latour and Woolgar (1986, 245) concluded that science is primarily concerned with the creation of an ordered account of reality, rather than the transfer of information pertaining to the world into human-understandable formats, and they characterize latter as the practices enacted in order to fulfill the former aim. However through a pragmatic lens, these two processes may be rendered as intertwined and inseparable threads of a unified experience of knowledge production.

Realizing this, Knorr Cetina (2001) highlighted how every action taken in the production of knowledge is underpinned by a desire to fill a gap in knowledge. Integral to this approach is her definition of knowledge objects, which are defined as settled statements about the world, which tend to be the goal of most scientific pursuits. Knorr-Cetina illustrates this as drawers in a filing cabinet, which have been organized squarely and that are accessible for retrieval. In contrast, gaps in knowledge or unexpected behaviour are characterized as untidy and incomplete drawers that exist in an unresolved state. Knowledge objects are comparable to Latour and Woolgar's (1986, 76) definition of facts as timeless and author-less accounts of reality, which have been elevated from the material systems that would otherwise constrain them as contingent to the specific material setup of the lab. However, Knorr-Cetina's concept renders knowledge production as a more dynamic, constructive and situated process that involves the use of already established knowledge in the validation of newly arrived at ideas (Knorr Cetina 2001, 181–82).

Her approach adds pragmatic and discursive elements, which help frame scholarly practices from the perspectives of the practitioners themselves, in terms of potentiality, certainty and desire. For instance, the desire of a lab technician to have the protein react as she expects it to will guide her practices, including her selection of reagents used to interact with it (Knorr Cetina 2001, 181). She thus alters the material assemblage of the system based upon her contextualization of what has or has not worked before, her educated hypothetical suppositions concerning the ways various actors might interact, and a reiteration of her goal. Her concern with the elusive nature of the protein thereby derives from her consideration of how the knowledge she might obtain will fit into the established order.

Knorr Cetina's work reveals how ANT does not adequately account for the circumstances through which the structures that support science come into being, nor the intentionality of human agents who assemble material apparatus to meet their goals. She refocuses attention on discursive aspects of knowledge production by considering expressions of potentiality, certainty and desire elicited by scientists as subjects acting with intent. This emphasis on the role of the subject, who conceives of other things from a unique perspective, introduces some asymmetry with respect to the roles of actors in the overall system. However, it does not totally dismiss the agency of non-human actors; it merely recognizes that the agency of objects are harnessed by humans and strategically configured into productive systems.

In other words, this approach recognizes that a scientific observer is needed to intervene, who ascribes meanings to materials, who delineates natural phenomena from modern cultural processes, and who selects signal from noise. It also enables science to be viewed as a collaborative and discursive process, comprising a convergence of methodologies and theoretical outlooks, often upon the same materials. This prompts recognition that 'the scientific record' exists in multiple forms, and varies according to the diverse range of perspectives that may characterize the objects of enquiry in different ways. When presenting their findings, scientists thus engage in discursive processes of semantic negotiation, comparing our own conceptions to those held by others.

Scientific data may therefore be considered as a connective medium that enables research to be conducted in a continuous and collaborative manner.

Continua of practice

This notion of a continuum or network of practice is characterized by recognition of the diverse array of work enacted through collective effort to contribute to information commons. It accomplishes this by taking a pragmatic approach to understanding the situatedness, historicity and social contingency of work (Dallas 2015, 440; 2016).⁴ Dallas (2016) is explicit in his promotion of a landscape metaphor whose ecosystems and inhabitants negotiate their differences in a pragmatic manner rather than one that emphasizes control and resource. Moreover, this relates to Cole's (1996) notion that culture is a permeable and layered substance that nurtures growth and development across uneven social landscapes. These views exemplify a broader trajectory that imagines knowledge work in ecological terms (cf. Star 1995).

In writing about specific issues regarding how scholarly content originating from various domains is professionally curated, Dallas brings attention to the ways in which such ecological metaphors might actually be applied. In his view, intellectual and pragmatic inquiry pertaining to the mindsets of the domains of knowledge production is necessary in order to better curate the meanings and intents behind archived material. Curation, for Dallas, is a continual process of interpretation and re-presentation; as objects (either physical or conceptual) are engaged with from various perspectives, they are ascribed new meanings. In other words, each interlocutor who engages with an object passing through their care as part of a curatorial process renews the object's meaning in ways that both reflect the object's history and shape its potential down the line. It is thus necessary, according to Dallas, to inquire about the ways in which domain specialists come to know an object and situate it as part of their lives in order to preserve and present the object in ways that adequately reflect its accrued meanings. This requires greater integration of domain specialists as part of curatorial workflows in order to better account for their outlooks as active stakeholders.

This outlook is interesting because it more clearly makes use of the notion that objects are sites of negotiation. Dallas (2015) applies these ideas to consider the ways in which the archaeological record is formulated and made meaningful in various ways throughout inherently collaborative, discursive and interdisciplinary archaeological projects. Indeed, his use-case is exemplary since relics, ruins and artefacts out of their original contexts serve as opportunistic lenses through which fragmentary glimpses of alternative realities might be captured. As Huvila (2016) notes, archaeological site reports, which are typically considered to represent the conclusive findings of a project, are produced from a series of complex negotiations among stakeholder specialists, but also present lopsided accounts of the knowledge recovered from the site; the organizational structure of archaeological projects, and the inherent epistemic limitations that guide the logistics and culture of the discipline, shape the final product and influence how it will be used in later secondary research contexts.

⁴It should be noted that Costis Dallas was my doctoral supervisor.

In this sense, archaeological projects resemble, to some degree, “wicked” problems, whose contradictions, incompleteness and social complexity render them resistant to completely satisfying resolutions (Rittel and Webber 1973). However, further inquiry into the ways in which meanings are ascribed throughout the course of a project, as they accrue via pragmatic and situated experiences, may help to gain a better understanding of the tensions and relationships at play. By highlighting the complex textures of work contributed by various members of a project, those who wish to use the generated knowledge in alternative arenas of practice may gain a better feel for the ways in which it was woven, and thus use it more effectively.

Paying close attention to the minute and pragmatic aspects of practice, is, according to Latour and Woolgar (1999, 24), what those in the fields of ‘science studies’ do best. By studying the particularities of what scientists actually do, and the ways in which they relate to things through series of coordinated social and technological mechanisms, we may be able to understand flows of knowledge across a continuum of understanding. By observing how professionals communicate and connect with each other and with the phenomena that captivate their interests, we can come to better understand the epistemic culture – the mindsets that drive curiosity and foster growth – which shapes the ways that they understand the world (Knorr Cetina 1999). As a result, scientific research, framed as a wide variety of active and cooperative engagements among people and objects, become situated alongside other ways of knowing. This is particularly important in a world where Mertonian notions of scholarly tribalism between the sciences, social sciences and humanities remain pervasive, despite many years of critique demonstrating that all scholarship actually occurs along an open continuum. In addition to helping foster more genuine cross-disciplinary encounters, recognition of the cultural situatedness of scholarship will enable scholars to more effectively operate in public and contribute towards the potential resolution of wicked problems.

- Blumer, Herbert. 1954. “What Is Wrong with Social Theory?” *American Sociological Review* 19 (1): 3–10. <https://doi.org/10.2307/2088165>.
- Bowen, Glenn A. 2006. “Grounded Theory and Sensitizing Concepts.” *International Journal of Qualitative Methods* 5 (3): 12–23. <https://doi.org/10.1177/160940690600500304>.
- Charmaz, Kathy. 2003. “Grounded Theory: Objectivist and Constructivist Methods.” In *Handbook of Qualitative Research*, edited by Norman K. Denzin and Yvonna S. Lincoln, 2nd ed., 249–91. Thousand Oaks, California: SAGE.
- . 2014. *Constructing Grounded Theory*. 2nd ed. SAGE.
- Cole, Michael. 1996. “Putting Culture in the Middle.” In *Cultural Psychology: A Once and Future Discipline*, 116–45. Belknap Press.
- Dallas, Costis. 2015. “Curating Archaeological Knowledge in the Digital Continuum: From Practice to Infrastructure.” *Open Archaeology* 1 (1): 176–207. <https://doi.org/10.1515/opar-2015-0011>.
- . 2016. “Digital Curation Beyond the ‘Wild Frontier’: A Pragmatic Approach.” *Archival Science* 16 (4): 421–57. <https://doi.org/10.1007/s10502-015-9252-6>.
- Haraway, Donna J. 1988. “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.” *Feminist Studies* 14 (3): 575–99. <https://doi.org/10.2307/3178066>.
- Huvila, Isto. 2016. “Awkwardness of Becoming a Boundary Object: Mangle and Materialities of Reports, Documentation Data, and the Archaeological Work.” *The Information Society* 32 (4): 280–97. <https://doi.org/10.1080/01972243.2016.1177763>.
- Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press. <https://books.google.com?id=g6nDQgAACAAJ>.
- . 2001. “Objectual Practice.” In *The Practice Turn in Contemporary Theory*, edited by Theodore R. Schatzki, Karin Knorr Cetina, and Eike von Savigny, 175–88. London; New York: Routledge.
- Latour, Bruno. 1992. “Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts.” In *Shaping Technology/Building Society: Studies in Sociotechnical Change*, edited by W. E. Bijker and John Law, 225–58. Cambridge, Massachusetts: MIT Press. <http://www.citeulike.org/group/718/article/3382023>.
- . 1999. *Pandora’s Hope: Essays on the Reality of Science Studies*. Harvard University Press. <https://books.google.com?id=RMu6wbzVrVkC>.
- Latour, Bruno, and Steve Woolgar. 1986. *Laboratory Life: The Construction of Scientific Facts*. 2nd ed. Princeton, New Jersey: Princeton University Press. <https://repository.library.georgetown.edu/handle/10822/811129>.

- Lave, Jean, and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge, United Kingdom: Cambridge University Press.
- Maryl, Maciej, Costis Dallas, Jennifer Edmond, Jessie Labov, Ingrida Kelpšienė, Michelle Doran, Marta Kołodziejska, and Klaudia Grabowska. 2020. “A Case Study Protocol for Meta-Research into Digital Practices in the Humanities.” *Digital Humanities Quarterly* 14 (3). <https://www.digitalhumanities.org/dhq/vol/14/3/000477/000477.html>.
- Nicolini, Davide. 2009. “Zooming In and Out: Studying Practices by Switching Theoretical Lenses and Trailing Connections.” *Organization Studies* 30 (12): 1391–1418. <https://doi.org/10.1177/0170840609349875>.
- Pickering, Andrew. 1992. “From Science as Knowledge to Science as Practice.” In *Science as Practice and Culture*, edited by Andrew Pickering, 1–26. University of Chicago Press.
- Rittel, H. W., and Melvin M. Webber. 1973. “Planning Problems Are Wicked.” *Polity* 4:155–69. http://cec.prodwebb.lu.se/sites/cec.prodwebb.lu.se/files/rittel_and_webber_1973_planning_problems_are_wicked_problems.pdf.
- Saldaña, Johnny. 2011. *Fundamentals of Qualitative Research*. Understanding Qualitative Research. New York: Oxford University Press.
- Selting, Margret, Peter Auer, and Dagmar Barth-Weingarten. 2011. “A System for Transcribing Talk-in-Interaction : GAT 2.” *Gesprächsforschung : Online-Zeitschrift Zur Verbalen Interaktion* 12:1–51. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/166>.
- Star, Susan Leigh. 1995. *Ecologies of Knowledge: Work and Politics in Science and Technology*. SUNY Press. <https://books.google.com?id=wpv9HZKaCnwC>.
- Suchman, Lucy. 2007. *Human-Machine Reconfigurations: Plans and Situated Actions*. 2nd ed. Cambridge University Press.
- VERBI Software. 2021. “MaxQDA 2022.” Berlin. <https://www.maxqda.com>.