Adapting to Challenges in Qualitative Fieldwork through Theoretical Sampling

Johanna Cohoon
hannah.cohoon@utah.edu
University of Utah
Salt Lake City, Utah, USA

James Howison
jhowison@ischool.utexas.edu
University of Texas at Austin
Austin, Texas, USA

ABSTRACT

We present a case study of theoretical sampling from a recent experience conducting grounded theory research. This example demonstrates how theoretical sampling can be used to adapt to challenges during data collection. By providing a detailed account of our methodological decisions and a description of the major aspects of data collection in a grounded theory study, this case study demystifies theoretical sampling. We provide several guiding questions that researchers can use to ensure thoughtful theoretical sampling. By committing to answering these questions, grounded theorists can preregister their work while maintaining analytical flexibility.

CCS CONCEPTS

• **Human-centered computing** → *Field studies*; *Computer supported cooperative work*.

KEYWORDS

grounded theory, theoretical sampling, open science, preregistration, research methods, trace data, case study

ACM Reference Format:
Johanna Cohoon and James Howison. 2023. Adapting to Challenges in Qualitative Fieldwork through Theoretical Sampling. In *CHI '23: CHI Conf. on Human Factors in Computing Systems Proceedings, April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3544549.3573873

1 INTRODUCTION

Grounded theory methods are an important tool for researchers in human computer interaction (HCI) and computer supported cooperative work (CSCW) to analyze their data. In this paper we present a case study in theoretical sampling within a ground theory project.

Grounded theory analyses have, for example, provided rich explanations of how virtual work is influenced by social worlds [6], how data scientists conceive of and interact with data [17], and how and why people with disabilities participate in digital crowdwork [22]. This methodological approach is useful for analyzing qualitative data without testing a priori hypotheses.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3573873

However, published descriptions of the execution of grounded theory studies are often shallow and abstract, with publications focusing on detail in the results and not in the method. This is ironic, given the affinity between grounded theory and deep empirical data and it also runs counter to the growing interest in scientific transparency through practices like preregistration[1]. Furthermore, it hampers readers' interpretation of published results and portrays grounded theory methods as infinitely flexible or mystified. Grounded theory methods do encompass several established traditions [2], contributing to the methodological haze.

The scientific community is experiencing a data deluge [13], providing fodder for qualitative studies of existing data. Public online content (e.g., digital traces and text corpora) is suited for grounded theory methods because the method allows for analysis of secondary data. As scholars develop new approaches like digital trace ethnography [9] that are often used in conjunction with grounded theory analysis, it is important to bring methodological detail into focus. The case presented here does just that.

Existing texts such as those handbooks written by the intellectual founders [11] and leaders [1–3, 15] of grounded theory provide detailed accounts of the method's execution. In this article, we provide a supplement to those comprehensive sources. We focus on one fundamental aspect of the grounded theory method, *theoretical sampling*, in order to show its utility to HCI researchers and provide a practical example of its execution. Theoretical sampling is considered "controlled opportunism" [5]—a chance to address gaps in the data by intentionally sampling for yet-unincorporated data. While similar, this method is distinct from what scholars scholars like Yin and Merriam refer to as purposive sampling, where data are chosen at the researcher's discretion but in advance, rather than throughout the course of data collection as is done in theoretical sampling [21]. Here we show how and why theoretical sampling can be implemented by leveraging a case from the authors' recent experience. This case effectively illustrates how theoretical sampling can help overcome practical issues during data collection.

Like Furniss et al. [8], our goal in presenting this case is not to expound the complete results of the executed study, but to provide methodological insight. In the following sections we first provide a brief overview of grounded theory methods and define some key features: constant comparison through coding and memoing, theoretical sampling, and theoretical saturation. Next, we summarize some key features of the Open Science Framework (OSF), the site for the research which provides the present examples of theoretical sampling. Given that context, we then describe how we enacted the

[1]Registration is the specification of data collection and analysis methods in a public setting such as an online repository or registry. When completed in advance of data collection, a registration is referred to as a preregistration. Analyses, recruitment plans, etc. may be included in a (pre)registration.

key features of grounded theory, focusing on an occasion where theoretical sampling helped overcome practical issues during data collection. Finally, we discuss the adaptability of theoretical sampling and pose guiding questions for theoretical sampling, noting that these may be useful for qualitative preregistration.

2 DATA COLLECTION IN A GROUNDED THEORY STUDY

Grounded theory analysis involves iteratively collecting and analyzing data to construct a theory that explains the observed phenomenon, and which may be generalized to other contexts. The iterative processes involved are referred to as *constant comparison* and *theoretical sampling*. These processes allow for an abductive approach to analysis [16] that produces a theory that is "integrated, consistent, plausible, close to the data" [11]. Rather than deduced via hypothesis testing, the theory is grounded in the data and internally validated such that it achieves *saturation*. This work may be informed by *sensitizing concepts* that provide "initial but tentative ideas to pursue" during analysis [1]. Engaging thoughtfully in these processes and acknowledging any sensitizing concepts promotes the reflexivity that most modern variations of grounded theory methods require [2].

2.1 Constant Comparison

Constant comparison is the fundamental process of grounded theory analysis; it refers to the analyst persistently making comparisons between data points in order to recognize patterns. Constant comparison is often carried out by writing analytical memos or coding data. Coding, in this sense, refers to the practice of labeling data in grounded theory analysis. "Through coding," writes Charmaz [1], "you define what is happening in the data and begin to grapple with what it means. [You begin making] generalizable theoretical statements that transcend specific times and places and contextual analyses of actions and events." Thus, researchers compare new data points to old data points, new labels to old labels, and continually refine the abstractions they produce via coding, memoing, and possibly other analytical techniques. Charmaz [1] instructs that statements within the same interview can be fruitfully compared, as can content across interviews and data sources. Constant comparison allows the researcher to develop and test the consistency of categories to represent aspects of their data, to integrate those categories and define them as concepts, and to delimit and make an explicit theory [3, 11].

2.2 Theoretical Sampling

Theoretical sampling can be summarized as collecting data because of insights that the analyst derived from other data. As new data are collected, they may prompt the researcher to pursue new directions or dig deeper into ongoing lines of analysis. Glaser and Strauss [11] indicate that theoretical sampling ensures that "data collection is *controlled* by the emerging theory." In this way theoretical sampling aids in delimiting the emergent theory; constant comparison ultimately yields a "smaller set of higher level concepts" [11] while theoretical sampling specifies a scope that can ensure "a close correspondence of theory and data" [11]. Theoretical sampling must be conducted with respect to the study's goals—the researcher cannot simply wander toward available data, they must seek out data that will inform their research questions and emergent findings. According to Glaser and Strauss [11], "The basic question in theoretical sampling. . . .is: *what* groups or subgroups does one turn to next in data collection? And for what theoretical purpose?" While the goal of theoretical sampling is to bring greater clarity to the constructed theory through the addition of more data for comparison, researchers are also advised to pursue lines of investigation they find exciting or surprising [1].

2.3 Saturation

Data saturation, also called theoretical saturation, is a standard criterion for arresting data collection and analysis in a grounded theory study. Despite its importance, recognizing saturation can be difficult. Charmaz [1] summarizes saturation as achieved when theoretical concepts are "robust because you have found no new properties" through comparison and "established properties [i.e., categories or concepts] account for patterns in your data."

To Charmaz, data collection does not end when novelty stops arriving with new data (this may never happen), but when the theoretical concepts that the researcher expressly chose to explore can withstand novel comparisons. Saturation therefore must be continuously evaluated during constant comparison and theoretical sampling; it requires sufficient data as well as sufficient interaction with that data [18]. Nelson [18] further suggests saturation is achieved through connection of the themes to prior literature and that it requires findings be reported with subtlety and nuance. When the relationships between saturated categories are explained, this forms a theory of the studied phenomenon.

3 METHODOLOGICAL SPECIFICITY IN ACCOUNTS OF GROUNDED THEORY

Grounded theory studies in HCI and CSCW may begin with secondary data or researchers may collect new data for their study [16]. Regardless of their origin, data analyzed with grounded theory methods must be iteratively assessed in the manner described above. However, despite the importance of constant comparison and theoretical sampling as procedures for grounded theory analysis, these steps are rarely described in detail in publications.

For example, Fitzpatrick, Kaplan, and Mansfield [6] present their highly cited ethnographic study of system admins' virtual work with little description of their methods. The authors forthrightly state their dependence on prior literature, and they declare their use of grounded theory methods, but the path from field notes to findings is obscured. How were early insights challenged as new data accrued?

Some authors like Muller et al. [17] acknowledge critical steps in grounded theory analysis like constant comparison and theoretical sampling. These researchers described their methods in studying data science workers:

"We returned to the passages, and to the inter-views, repeatedly as necessary (constant comparison), to look for additional evidence and to test and revise our emer-gent understanding (theoretical sampling, abductive logic). These iterative analyses led to a core set of 19 axial codes, which we combined into the 5 selective codes of section 4."

Yet, while their sample is well explained, reflection on their strategy for accruing new data in light of old data is absent. How did those 19 codes or their combination affect the semi-structured interviews that were held in the later stages of data collection? Only the quantitative summary quoted here is given.

Other authors like Zyskowski et al. [22] also identify important procedures of grounded theory analysis and go further by describing an approach to theoretical sampling: they developed a survey based on the emergent themes from interviews. However, what those emergent themes were and how they affected the survey design is not included in the paper.

The limited space allotted to publications is prohibitive of the kind of rich methodological explanations necessary to completely describe the iterative and data-heavy nature of grounded theory analysis. Nevertheless, without examples of the kinds of comparisons made between data points or summaries of analytical strategies, learning about the execution of grounded theory methods remains difficult. Readers must be willing to trust authors' reasoning and novices are provided few examples from which to learn. Muller and Kogan [16] argue that more writing on grounded theory methods in HCI and CSCW is necessary to advance our application of the tools—to learn "what works and what does not."

With the case presented below, we show what worked for a study attempting to leverage trace data to show routine use of an open science system. We describe several instances of theoretical sampling and show that, in addition to its utility for generating insights, the flexibility of the grounded theory procedure was useful for overcoming roadblocks encountered during trace data collection. We explain methodological decisions in detail to demystify the execution of grounded theory for HCI researchers.

4 A CASE FOR THEORETICAL SAMPLING

To provide concrete examples of theoretical sampling for other researchers and to illustrate the utility of the approach, we describe how theoretical sampling was executed during a qualitative study of OSF engagement. We begin with an overview of OSF's features before presenting an account of our progression through the components of grounded theory analysis defined above. This account focuses on how theoretical sampling overcame a challenge encountered during data collection.

It should be noted that as a case study drawn from a larger project the methods discussed here are not complete descriptions of the work completed. However, the examples provided in this section are representative and accurately portray our use of theoretical sampling. Table 1 summarizes the components of grounded theory we enacted and their effects on our data collection.

4.1 OSF

OSF (osf.io) is a web application for researchers to share and document their work online. Users of OSF can create projects to organize their research. Projects are made up of components that have most of the same features as their parent projects: a title, a contributor list, a wiki, file storage, tags, a unique URL, and timestamped activity logs (see Figure 1). On OSF, projects and components can be arranged hierarchically and can link to other projects, preprints, registrations, or publications. Users can collaborate with one another on projects and components of those projects.

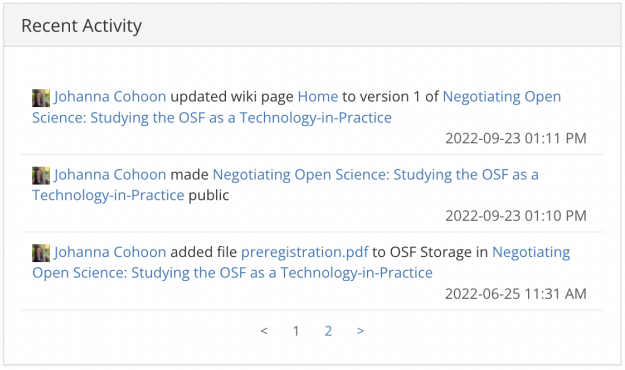

Figure 1: One author's Recent Activity from their OSF project

OSF includes a registration feature to show the status of a project at a given point in time. When a project is registered before data collection begins, this feature is useful for demonstrating at the conclusion of a project that all a priori hypotheses have been tested as planned. When registering a project on OSF, users click a button and are guided through a step-by-step form that asks them to disclose facts about their research study. These responses are associated in the OSF database with a time-stamped version of the source project that shows its state at the time of registration. Viewers of a public registration can see the form responses and archival version of the project. Because they are distinct objects in OSF's database, public registrations can be retrieved programmatically through the OSF API.

4.2 Sensitizing Concepts

During the project a literature review was conducted to better understand the influence of scientific infrastructure on research practice and engagement with open science. This review and a theoretical framework provided the sensitizing concepts that shaped initial data collection.

In sum, this review showed that technology such as OSF can *persuade users through design* [7]. However, it also indicated the possibility of *value conflicts* between system developers and users [4] as well as other unanticipated forms of engagement with open science technology. *Open science practices* like the data sharing and preregistration that occur on OSF were shown to be *increasing* across scientific disciplines [20]. However, many researchers still have concerns about engaging in open science [14].

The theoretical framework included *structuration theory* [10] and *technologies-in-practice* [19]. This framework encouraged the consideration of how human agents *routinely* draw on their material and societal contexts as *rules* or *resources* for affecting change. This structuration is recursive such that rules or resources from one scenario may become rules or resources in the next.

Based on the literature review and theoretical framework, the research focus was more clearly delineated. More than a study of

Component of GT Analysis	Action Taken	Examples of Effect on Data
Sensitizing concepts	Conducted a literature review	Chose methods to capture routine engagement with OSF and the factors that affect that engagement, recognizing that different populations (users, non-users, developers) may have different concerns. Labeled data with codes relevant to/inspired by sensitizing concepts
Constant comparison	Compared codes and coded data	Re-coded data, created new codes, and provisionally accepted codes' adequacy
Constant comparison	Compared sources of data	Pursued a theoretical sample: Recruited from populations poorly represented in the data (e.g., non-users or people who had registered research)
Theoretical sampling	Collecting (and comparing) an abundance of data relevant to emergent codes of interest	Redesigned interview protocols to capture more data relevant to emergent codes (e.g., asking about adding contributors to digitally shared projects)
Constant comparison and establishing saturation	Triangulated results through comparison of multiple data types	Pursued a theoretical sample: Sought multiple forms of data to challenge/support emergent ideas (e.g., self-report data through interviews and behavioral data through traces). When one collection strategy (web scraping for traces of behavioral data) failed, pursued other data that could also challenge/support emergent ideas

Table 1: Summary of examples from our execution of grounded theory methods

OSF engagement, it would explore the enactment of open science through persuasive technology. We anticipated that OSF developers, users, and non-users might *routinely* engage with the platform in different ways from one another, possibly because of differing *values*. While no hypotheses were made, the data collection strategy was designed to detect such differences in values and practices. Interview protocols, for example, included discussion of scientific values. Observation was planned to capture evidence of developers' routine behaviors and trace data and documents were similarly desired to provide evidence of participants' engagement with OSF. The multiple forms of data would enable the triangulation of findings. Trace data in particular were useful because they are artifacts of system use, providing evidence of both user behavior and system design [12, 23].

While data was collected and coded, the sensitizing concepts inspired some codes. For example, the use of *resources* was often coded with a combination of "leveraging" and the resource itself, e.g., LEVERAGING: Github.

4.3 Data Collection and Analysis

As data accumulated, it underwent constant comparison. One approach to comparison was to simply view the data in ATLAS.ti, selecting a particular code or interview and comparing statements. Figure 2 shows how ATLAS.ti lists the data that was labeled with the code "LEVERAGING: Github" which was introduced to identify a resource that study participants took advantage of. Items previewed in these lists could be displayed in full. By reviewing the data in lists like these, it was easy to consider the differences and similarities between cases as well as how well the codes fit the data. An outcome of this form of constant comparison might be re-coded data, a new code made from the combination or splitting of others, or an acceptance that first instincts were (provisionally) correct.

Comparison also made salient where recruitment efforts needed to be directed, answering Glaser and Strauss's (1967) question of where to turn next for data collection. For example, comparison of data showed that the data was not sufficiently representative of non-users—especially those completely unfamiliar with OSF. Thus, recruitment efforts were re-focused on that population. In this way, constant comparison directed theoretical sampling. Theoretical sampling also prompted recruitment of participants who had registered a study via OSF because, while developer interviews had covered registrations, there was little in-depth data on users' experiences with registration.

Theoretical sampling not only directed recruitment, but also the design of interview protocols. The technique was used to ensure that interview questions would shed light on the emergent codes perceived most important. For instance, interviews and trace data showed that contributor lists on OSF projects were not always reflective of the full research team. On other occasions the list of names might be complete but not all of the contributors listed had

Figure 2: ATLAS.ti Displaying data coded with "LEVERAGING: Github"

OSF accounts. These data were labeled with codes like "adding contributors" and "ALIGNING: contributor lists." When the opportunity arose to ask a non-OSF-user about a similar behavior, it was taken: When N1 showed how they stored their working files in Box, they were asked about who they shared those files with. Thinking of those researchers with un-accepted invitations to be contributors on an OSF project, the first author asked N1, "Have you ever had to prompt someone to accept one of those invitations [to a Box folder]?" This allowed comparison of OSF use to similar activity on other platforms, yielding greater understanding of how researchers manage contributor lists and permissions. Here, theoretical sampling was enacted through an interview question that solicited material for constant comparison.

These examples show that constant comparison prompted theoretical sampling: we sought more users who had registered research and more non-users. As coding and memoing progressed, emergent patterns were identified and those informed the ongoing data collection strategy.

4.4 Challenges with Trace Data

Among the themes that were treated with analytical priority was OSF developers' redesign of the registration feature to better meet user expectations. Another was the organization of materials and information through OSF projects. These and other analytical avenues were documented in memos and the codes applied to existing data. When we met challenges during data collection, we strategized by collecting a theoretical sample that could speak to these themes. Because theoretical sampling is conducted based on emergent findings rather than *a priori* hypotheses, this approach allowed us to adapt to the challenges we faced.

Originally, we agreed that the logs from users' Recent Activity panels on OSF would be useful traces to gather. However, after multiple failed attempts to gather these traces through web scraping and API tools, a new strategy was needed. Both authors met to discuss the issue. We acknowledged that analysis of other data was fairly advanced—major themes were emerging. Meanwhile, the plan for analyzing traces from Recent Activity panels was unclear and untied to those developing findings.

Early in the project we had discussed quantifying things like the most common and most recent actions taken by OSF users. Those findings, we reasoned at the outset of the research, might prompt pursuit of new analytical directions. Yet, we reasoned months into data collection, constant comparison had already provided ripe analytical avenues and turning away from them because frequency statistics could be available made little sense. Deciding that capturing the activity logs would involve too much additional effort and provide uncertain returns, the authors agreed that gaining further clarity on the organizational schemes of OSF projects was a better use of time. That data, while not traces gathered through the Recent Activity panel as planned, would meet the needs of the theoretical sample. In light of that conversation, the first author began mapping the structure of participants' OSF projects.

4.5 Project Mapping

Mapping OSF projects involved using Miro (miro.com) to diagram the first project listed in each user's OSF profile. Maps showed the hierarchy of the project and its components and any other objects that were linked to (e.g., a registration or a preprint). During this process, the first author started noticing that users sometimes uploaded a file that they labeled as a preregistration. Thinking this

should be considered further in a memo, she documented those files in addition to the registrations created through the OSF interface. When reviewing the completed project maps, it became clear that both forms of registration were commonly used.

Project maps supplanted Recent Activity trace data effectively because they spoke to themes of interest like registration and organization of OSF projects. They also supported the triangulation of findings by providing an additional, behavioral data source. These examples of real-world use could be compared to developers' descriptions of how OSF is meant to be used.

4.6 Gaining Insights

Theoretical sampling directed data collection such that concepts and practices that emerged as important could be rendered in higher resolution. For example, the OSF project maps rendered in greater clarity some key differences in OSF developers' and users' understanding of registrations. Specifically, some users of OSF considered their research to be registered when they uploaded files explaining their study methods in advance of its execution. However, OSF developers consider registration to be a native feature of the platform that involves creating an archived version of the user's project (a citable container for files and wiki text). Only this latter form of registration can be retrieved through an OSF API call—a fact that had multiple implications for the study results. These findings were evident through interviews, but project maps showed the extent of users' confusion—some users leveraged the native feature, some users uploaded files, some users did both within the same project. Even if registrations were not a dominant topic during a given interview, these maps provided evidence of a user's understanding of registering research.

Project maps were an outcome of theoretical sampling and proved vital to establishing saturation. Through comparison with the data from interviews, observation, traces, and documents, these project maps helped bring nuance and depth to emergent findings. The multiple sources of data triangulated the results by providing internal validity.

While finalizing construction of the grounded theory (of the enactment of open science via persuasive technology), the insight about the multiple interpretations of registration on OSF served as an important connecting thread between concepts. It proved to be a key example in the study's findings.

5 DISCUSSION

In this case study, we have described how we enacted theoretical sampling and other grounded theory techniques to gather multiple forms of qualitative data for a study of OSF engagement. The major findings of the source study heavily referenced the insight that OSF developers and users conceptualize research registration differently from one another. While it resonated with the sensitizing concepts established through literature review, the insight was not anticipated or hypothesized. Instead, it was cultivated through a series of methodological decisions. Although this case study reported those decisions in detail, that kind of description is often absent from grounded theory publications.

With detailed descriptions of grounded theory methods largely unavailable in empirical literature, newcomers to the method have few examples to learn from. Readers are presented with accounts of grounded theory that gloss over its complexity rather than ones that teach "what works and what does not" [16]. This case study hopes to counter that trend, providing multiple examples of the decision making that occurs during the execution of grounded theory studies. Through the experience reported here, we have found it important for grounded theory research to conscientiously embrace the adaptability afforded by theoretical sampling.

By adaptability, we refer to the fact that theoretical sampling asks for thoughtful pursuit of informative data rather than dogmatic adherence to a specific procedure. In our experience, this characteristic of the technique proved helpful when we failed multiple times to programmatically collect trace data about user behavior. This challenge recalls Furniss et al.'s [8] point that grounded theory studies must be "managed alongside their practical constraints." By reflecting on the theoretical purpose of the traces and considering the emergent analytical outcomes, we were able to adapt our approach and identify a new data source that was informative and fit our practical constraints.

Notably, the sampling needs changed as data collection proceeded. At one point, we collected a theoretical sample by recruiting more non-users. At another, by accruing more examples of managing contributors to digital projects. At another, by collecting a different form of data altogether. Because our sampling strategy responded to the insights derived from earlier data, that strategy evolved along with our analyses. Again, this dynamism is an inherent quality of theoretical sampling that allowed us to overcome the challenge of trace data collection.

Theoretical sampling is not a license to change tack without cause or acknowledgment. Researchers seeking to harness the adaptability of the grounded theory technique should do so conscientiously. We suggest several questions to answer for grounded theorists identifying a theoretical sample:

- Why are these data the right ones for the study?
- How will these data speak to the research questions, sensitizing concepts, or emergent findings?
- Which findings prompted changes in recruitment strategies?
- What aspects of recruitment did theoretical sampling influence?
- What ways did it impact research materials like interview protocols or collected traces?

Space restrictions may prohibit answering these questions in the body of an article, but they are excellent topics for memoing and could be included as a publication supplement so that others can learn from the work done by the grounded theorist. By answering these questions, authors would recount the key details necessary for readers to understand their sampling strategy and its impact on their findings.

Finally, qualitative research is often thought of as poorly aligned with transparency practices such as preregistration. Because qualitative findings like those in grounded theory studies are often constructed rather than experimentally discerned, researchers find it difficult to describe their process in advance. Preregistration forms asking for hypotheses demonstrate this positivist outlook–one that is contrary to much qualitative work. However, by publicly

committing in advance to answering questions like those above, researchers can promise post-hoc transparency and make themselves accountable to critically engaging with their data. Preregistration in this manner might lead to fewer mystified accounts of grounded theory studies while allowing needed flexibility.

6 CONCLUSION

The case study reported here served to demystify theoretical sampling in grounded theory research. Furthermore, we showed the adaptability afforded by the technique to be useful for overcoming challenges during data collection. Based on our experience, we propose some questions to guide theoretical sampling in future studies. Answering these questions can aid analysis as well as help others understand the warrant for claims. While those answers may be revealed at the conclusion of a study, committing to providing them via preregistration may still improve research quality and transparency.

ACKNOWLEDGMENTS

Our sincere thanks go to the participants in the research described above including Aliakbar Akbaritabar; Bilal Hayat Butt (Pakistan); Simon Columbus; Kellen Parker van Dam; Hendrik Erz; Brian J. Geiger; Florian Goth (SFB1170, Project Z03); Pen-Yuan Hsing; Lucy Ko; Chris C. Martin; Spencer Mermelstein, PhD; Brian Nosek; Gobinath Ravindran; Breno Souza-Marques; Andy Tseng; and Luiz Victorino. Thanks also go to Sarah Elaine Bratt and the Jason Weise's 2022 Human Centered Computing course for their feedback on drafts of this article.

REFERENCES

[1] Kathy Charmaz. 2014. *Constructing Grounded Theory* (2 ed.). SAGE.

[2] Adele Clarke, Carrie Friese, and Rachel S. Washburn. 2018. *Situational Analysis: Grounded Theory After the Interpretive Turn* (2 ed.). SAGE Publications, Los Angeles, CA, USA.

[3] Juliet Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology* 13, 1 (1990), 19.

[4] Janet Davis and Lisa Nathan. 2015. Value Sensitive Design: Applications, Adaptations, and Critiques. In *Handbook of ethics, values, and technological design: sources, theory, values and application domains*, Jeroen van den Hoven (Ed.). Springer, Dordrecht, 11–40. OCLC: 925374330.

[5] Kathleen M. Eisenhardt. 1989. Building Theories from Case Study Research. *The Academy of Management Review* 14, 4 (Oct. 1989), 532. https://doi.org/10.2307/258557

[6] Geraldine Fitzpatrick, Simon Kaplan, and Tim Mansfield. 1996. Physical Spaces, Virtual Places and Social Worlds: A study work in the virtual. In *Computer Supported Cooperative Work '96*. ACM, Cambridge, MA, USA, 10.

[7] B.J. Fogg. 2003. *Persuasive Technology: Using Computers to Change What We Think and Do*. Elsevier, Burlington. http://www.UTXA.eblib.com/EBLWeb/patron?target=patron&extendedid=P2943030& OCLC: 476057812.

[8] Dominic Furniss, Ann Blandford, and Paul Curzon. 2011. Confessions from a grounded theory PhD: experiences and lessons learnt. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 113–122. https://doi.org/10.1145/1978942.1978960

[9] R Stuart Geiger and David Ribes. 2010. The work of sustaining order in wikipedia: the banning of a vandal. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. ACM, Savannah, Georgia, USA, 10.

[10] Anthony Giddens. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press. Google-Books-ID: x2bf4g9Z6ZwC.

[11] Barney G. Glaser and Anselm L. Strauss. 1967. *The discovery of grounded theory: strategies for qualitative research*. Aldine Pub. Co., Chicago. http://hdl.handle.net/2027/hvd.32044018542340

[12] Sean Goggins, Christopher Mascaro, and Giuseppe Valetto. 2013. Group informatics: A methodological approach and ontology for sociotechnical group research. *Journal of the American Society for Information Science and Technology* 64, 3 (March 2013), 516–539. https://doi.org/10.1002/asi.22802

[13] Tony Hey and Anne Trefethen. 2003. The Data Deluge: An e-Science Perspective. In *Wiley Series in Communications Networking & Distributed Systems*, Fran Berman, Geoffrey Fox, and Tony Hey (Eds.). John Wiley & Sons, Ltd, Chichester, UK, 809–824. https://doi.org/10.1002/0470867167.ch36

[14] Nadine Levin and Sabina Leonelli. 2017. How Does One "Open" Science? Questions of Value in Biological Research. *Science, Technology, & Human Values* 42, 2 (March 2017), 280–305. https://doi.org/10.1177/0162243916672071

[15] Janice Morse, Barbara Bowers, Kathy Charmaz, Adele Clarke, Juliet Corbin, and Phyllis Stern. 2021. *Developing Grounded Theory: The Second Generation*. Rutledge.

[16] Michael Muller and Sandra Kogan. 2012. Grounded Theory Method in Human–Computer Interaction and ComputerSupported Cooperative Work. In *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition*. Taylor & Francis Group, Baton Rouge, UNITED STATES. http://ebookcentral.proquest.com/lib/utah/detail.action?docID=911990

[17] Michael Muller, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q. Vera Liao, Casey Dugan, and Thomas Erickson. 2019. How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–15. https://doi.org/10.1145/3290605.3300356

[18] James Nelson. 2017. Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. *Qualitative Research* 17, 5 (Oct. 2017), 554–570. https://doi.org/10.1177/1468794116679873 Publisher: SAGE Publications.

[19] Wanda J. Orlikowski. 2000. Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. *Organization Science* 11, 4 (2000), 404–428. http://www.jstor.org/stable/2640412

[20] Carol Tenopir, Natalie M. Rice, Suzie Allard, Lynn Baird, Josh Borycz, Lisa Christian, Bruce Grant, Robert Olendorf, and Robert J. Sandusky. 2020. Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLOS ONE* 15, 3 (March 2020), e0229003. https://doi.org/10.1371/journal.pone.0229003

[21] Bedrettin Yazan. 2015. Three Approaches to Case Study Methods in Education: Yin, Merriam, and Stake. *The Qualitative Report* 20, 2 (Feb. 2015), 134–152. https://doi.org/10.46743/2160-3715/2015.2102

[22] Kathryn Zyskowski, Meredith Ringel Morris, Jeffrey P. Bigham, Mary L. Gray, and Shaun K. Kane. 2015. Accessible Crowdwork?: Understanding the Value in and Challenge of Microtask Employment for People with Disabilities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, Vancouver BC Canada, 1682–1693. https://doi.org/10.1145/2675133.2675158

[23] Carsten Østerlund, Kevin Crowston, and Corey Jackson. 2020. Building an Apparatus: Refractive, Reflective, and Diffractive Readings of Trace Data. *Journal of the Association for Information Systems* 20, 1 (2020), 1–22. https://doi.org/10.17705/1jais.00590