Journal of the Association for Information Systems

Volume 27 | Issue 1 Article 10

2026

Theorizing From Contexts in Research With Digital Trace Data

Emmanuelle Vaast
McGill University, emmanuelle.vaast@mcgill.ca

Recommended Citation
Vaast, Emmanuelle (2026) "Theorizing From Contexts in Research With Digital Trace Data," *Journal of the Association for Information Systems*, 27(1), 1-18.
DOI: 10.17705/1jais.00955
Available at: https://aisel.aisnet.org/jais/vol27/iss1/10

Journal of the Association for Information Systems (2026) **27**(1), 1-18
doi: 10.17705/1jais.00955

EDITORIAL

ISSN 1536-9323

Theorizing From Contexts in Research With Digital Trace Data

Emmanuelle Vaast[1]
[1]Desautels Faculty of Management, McGill University, Canada, emmanuelle.vaast@mcgill.ca

Abstract

Many researchers rely on digital trace data, which are organically produced by sociotechnical systems and can provide insights into new digitally related phenomena. Yet theorizing from contexts observed through digital trace data is challenging. Digital trace data are part of these contexts. For some, these contexts may be long-established but changing, while for others, they may be new but perhaps not fully unprecedented. Theorizing from contexts observed through digital trace data introduces a dilemma between staying true to these contexts and developing new theory, which can overlook key aspects of the contexts or overstate how distinctive the observations really are. Theorizing from contexts with digital trace data thus involves developing new theory that both makes sense of and transcends these contexts. This study elaborates on the challenges and opportunities associated with theorizing from contexts with digital trace data and provides guidance on how to do so. It explains how researchers can engage in probing and elucidating contexts when analyzing digital trace data. Probing contexts involves surfacing the omnibus context and identifying, scanning, and connecting the discrete contexts from which digital trace data originate. Elucidating contexts involves situating, depicting, and explaining contexts to answer contextual *where, when, what, who, how,* and *why* questions. The progression of these elucidating questions can help researchers build theory from contexts observed through digital trace data. This study illustrates this framework using three papers that rely on digital trace data and examine distinct contexts. This framework can help researchers deepen their engagement with contexts as they analyze digital trace data and provide inspiration to build new theory.

Keywords: Digital Trace Data, Contexts, Computationally Intensive Theory Construction, Contextualizing, Theory Building, Theorizing, Phenomenon, Omnibus/Discrete Contexts

Hani Safadi was the accepting senior editor. This editorial was submitted on November 23, 2024, and underwent two revisions.

1 Introduction

Digital trace data have become an increasingly important source of data for researchers in multiple disciplines, including information systems (IS) (Beer & Burrows, 2007; Hartl et al., 2023; Lazer et al., 2009; Rogers, 2013). Digital trace data correspond to "records of activity (trace data) undertaken through an online information system (thus, digital). A trace is a mark left as a sign of passage; it is recorded evidence that something has occurred in the past" (Howison et al., 2011, p. 769). A number of researchers have considered the benefits and challenges of relying on digital trace data in their investigations (boyd & Crawford, 2011; Freelon, 2014; Grover et al., 2020; Østerlund et al., 2020; Vial, 2019).

Digital trace data are produced organically in particular contexts (Ash et al., 2024; Revina et al., 2023) and can be collected and analyzed by researchers. An increasing body of research has focused on how to theorize with digital trace data through, for example, explaining the different stages of research integrating human and computational analysis (Lindberg, 2020), elaborating new methods dedicated to the construction of computationally intensive theory (Berente et al., 2019; Miranda et al., 2022a), process and temporal theorizing (Kishore et al., 2024; Pentland et al., 2021), and discussions on the respective roles of theory building and the development of research as a platform with digital trace data (Grisold et al., 2023; Grover & Lyytinen, 2023).

Further, research on digital trace data is increasingly recognizing the importance of context (Grisold et al., 2024b; Pentland et al., 2020; Revina et al., 2023). A context constitutes the specific setting in which a particular situation, process, or, more generally, a phenomenon, takes place (Johns, 2006; Whetten, 2009). Considering context is essential to theorizing and to the research process (Dyer & Wilkins, 1991; Hong et al., 2014; Johns, 2006). Researchers have shown how explicitly considering context can reveal unexpected insights (Pentland et al., 2020) and aid in the development of "impactful situated explanations" (Grisold et al., 2024b).

In contexts associated with digital trace data, interesting phenomena appear and unfold (Vicari & Kirby, 2023). Contexts underlie and help researchers make sense of situations and processes observed through digital trace data (Grisold et al., 2024b). In response to the call to theorize phenomena (Fisher et al., 2021), especially digitally related phenomena (Monteiro et al., 2022), questioning context can help researchers connect digital trace data to phenomenon-based theorizing. This also relates to calls for a "push to the edges" (Grover & Lyytinen, 2015) of research, owing to a greater and deeper understanding of context. Researchers engaged in such ventures seek to deepen their grasp of context while still aiming to build theory (Grover & Lyytinen, 2023; Vaast & Walsham, 2013), which can be challenging.

The dilemma between contextualizing and theorizing, i.e., between developing a thick and specific understanding of one or more contexts and abstracting from these contexts to build theory that transcends them, is not new to the scholarship (Davison & Martinsons, 2016; Dyer & Wilkins, 1991; Whetten, 2009). However, for research relying on digital trace data, there are new aspects to this dilemma that make it distinctive and critically important (Grover & Lyytinen, 2023; Pousti et al., 2021). For one thing, digital trace data involve "context shifts" that are complex but also require specific theorizing (Baiyere et al., 2023). Further, while existing digital trace data scholarship can illuminate familiar contexts such as organizations (Avital et al., 2023; Grisold et al., 2024b; Pentland et al., 2021), digital trace data may also be associated with multiple other contexts (e.g., Zhang et al., 2020; Zhang et al., 2022) that researchers need to examine, and theorizing from these other contexts requires that researchers first develop an understanding of them. Also, there may be new, digitally enabled contexts such as social media-based contexts (Johns, 2018; Vaast & Pinsonneault, 2022) that feature new actors (e.g., bots) (Salge et al., 2022). Finally, digital trace data may be associated with multiple contexts, and there may be a need to disentangle the connections between online and offline contexts in research relying on digital trace data (Tarafdar & Kajal Ray, 2021; Vaast et al., 2017).

Theorizing from contexts accessed through digital trace data is thus an important yet challenging endeavor. The purpose of this paper is to provide researchers with guidance on how to theorize from contexts with digital trace data. If "context is king" (Davison & Martinsons, 2016) and needs to be explicitly considered in scholarship, research relying on digital trace data will require special effort in questioning the contexts that relate to the data. This study offers a framework that articulates how to probe and then elucidate contexts for research relying on digital trace data. Probing contexts involves researchers engaging in surfacing the broader omnibus context and then identifying, scanning, and connecting the discrete contexts that are the focal loci of their investigations. Elucidating contexts complements probing, as researchers building theory must situate, depict, and explain contexts. Elucidating contexts involves researchers asking contextual *where, when, what, who, how,* and *why* questions. After discussing the framework, this paper then illustrates it using three papers that leverage digital trace data from distinct contexts and build theory.

2 Contexts and Theorizing From Contexts in Digital Trace Data Research

2.1 Contexts and Their Importance for Theorizing

"In abstract terms, context is the set of circumstances in which phenomena (e.g., entities) are situated." (Griffin, 2007, p. 860). The phenomena that researchers investigate and problematize (Monteiro et al., 2022) are situated within contexts. Contexts shape why phenomena appear and evolve the way they do (Johns, 2018).

Contexts are important in theory and theory building because they shape the opportunities and constraints that participants experience as well as the conditions for processes to unfold (Johns, 2006, 2018). The meanings

of people's actions and behaviors are context dependent (Bamberger, 2008). Scholarship in international business, for instance, must contend with culture- and country-specific meanings and consequences (Welch et al., 2022). Behaviors that appear similar on the surface may have different meanings or consequences, depending on their cultural or organizational context (Davison & Martinsons, 2016). Apprehending context allows researchers to uncover the motivations and rationales behind actions and is thus essential for making sense of how and why phenomena occur. Therefore, context is a key explanatory tool rather than a simply descriptive material for theory building (Hong et al., 2014; Welch et al., 2022).

Context is particularly important for midrange theories (in contrast to grand theories) (Merton, 1957). Many social scientists, including IS researchers, aspire to build midrange theories, which operate between highly abstract, broad theoretical frameworks and narrowly focused empirical observations or hypotheses. Midrange theories incorporate context, and their applicability and mechanisms depend on contextual conditions. Unlike grand theories that aim for universal applicability, midrange theories typically focus on how certain dynamics operate in contexts.

Defining and delimiting the boundaries of investigated contexts helps researchers delineate the boundaries of a midrange theory's applicability and provide insights into when, where, and why certain phenomena occur. Johns (2006) notably distinguished between omnibus and discrete contexts. The omnibus context is broad and general, providing a backdrop that contributes to shaping behaviors or events across multiple situations. Discrete contexts are narrow and situational and contribute more immediately to shaping behaviors and events in a particular instance or setting.

Context matters in theory building because of the value inherent in rich, contextually grounded insights compared to abstract constructs (Dyer & Wilkins, 1991). Midrange theorizing should be deeply embedded in and derived from investigated contexts. It relies on deep rather than surface-level descriptions and on explanations that are intricately connected to contexts.

2.2 Theorizing From Contexts With Digital Trace Data

Theorizing from contexts is, however, challenging. Developing a thick rather than thin understanding of contexts is essential to understanding what is being investigated, but it can make it difficult to theorize in a way that will be applicable beyond the observed contexts. Researchers thus face a challenge in that the deeper their understanding of contexts, the more specific the theory they will build. This limits the applicability of their work to different contexts.

This issue is heightened when contexts are observed and analyzed through digital trace data. Putting digital trace data into their contexts and theorizing from the contexts of these digital trace data can be difficult for researchers (Breiter & Hepp, 2018; Hampton, 2017; Pousti et al., 2021). Moreover, digital trace data can emerge from and represent digitally enabled contexts that may be novel and underexplored (Johns, 2018). Thus, as they collect and analyze digital trace data, researchers may face the opposite perils of overstating the new and unprecedented and overlooking long-standing theories that would fit the contexts or underestimating what is original and distinctive about the contexts and theorizing too broadly about them.

Whetten (2009) provided insights that originated from the challenges of theorizing across different national contexts but can be expanded to account for the challenges and opportunities of digital trace data. Whetten examined the interface between theory and context and considered researchers' aim to develop theories *in* context (see Hong et al., 2014; Venkatesh, 2025) and/or theories *of* context (the focus of this study, what I call theorizing from contexts).

With theories *in* context, digital trace data are considered indicative of and even part of the investigated context. Researchers can analyze these digital trace data to make sense of a phenomenon and refine existing theory that is sensitive to the examined context (e.g., Da Cunha & Orlikowski, 2008; Xu & Zhang, 2022). Theories in context involve refining existing theory with digital trace data to make sense of a particular context (Choi, 2020). Researchers thus rely on existing theories to analyze digital trace data. They then adjust the existing theory to fit and respect the examined contexts and what they reveal. This approach explains patterns observed in digital trace data in light of existing theories, refining concepts and their relations to account for the contextual conditions unearthed through digital trace data.

With theorizing *from* contexts, digital trace data are viewed as more closely connected to the contexts and as even potentially constituting the contexts. This can happen when participants in the contexts act and interact entirely or partly online (e.g., online communities, open-source communities, social media). As they theorize from contexts, researchers aim at developing new theory that illuminates the contexts observed through digital trace data and the focal phenomenon.

While both approaches (theories in context and theorizing from contexts) are worthwhile, this paper focuses on theorizing from contexts. The reason for this is that for theories in context, it is more appropriate to consider context when testing or extending the theory (Jiang et al., 2022). In contrast, theorizing from contexts with digital

trace data is more conducive to developing new theory associated with digital trace data and to making sense of digitally related phenomena (Grover & Lyytinen, 2023). Theorizing from contexts involves developing new theoretical insights specifically around the contextual conditions that digital trace data reveal. Rather than applying existing theories to interpret data, theorizing from contexts involves relying on digital trace data to define and delve into what the contexts actually mean and reveal (Grisold et al., 2024b). Theorizing from contexts can illuminate emergent and uniquely digital—or digital-first—phenomena (Baskerville et al., 2020). With digital trace data, researchers can uncover new contexts and phenomena that appear online (Grisold et al., 2023), such as bot behavior, gig work, and algorithmic management (Shevchuk et al., 2021). Theorizing from contexts enables researchers to make sense of digital phenomena that may be new and may not be observable with data other than digital trace data (Pentland et al., 2021).

3 Challenges and Opportunities of Digital Trace Data to Theorize From Contexts

Theorizing from contexts observed through digital trace data can, however, be challenging for researchers due to the characteristics of digital trace data (Pousti et al., 2021) and the complex ways in which contexts manifest themselves and shift online (Vaast & Pinsonneault, 2022). Accessing digital trace data does not guarantee that researchers will gain the deep understanding of contexts that is critical for theorizing. This can make digital trace data research seem acontextual, i.e., lacking context and circumstances (Freelon, 2014). Yet digital trace data also offer opportunities for researchers to theorize from contexts (Salganik, 2019). This section elaborates on the challenges and opportunities for theorizing from the contexts that digital trace data offer researchers.

3.1 Challenges

3.1.1 Context Collapse

Context collapse, which leads to researchers being unable to understand the context of investigation, is particularly notable with digital trace data (Marwick & boyd, 2010; Pike et al., 2018). Context collapse refers to the "flattening" of different participants and contexts into digital trace data (Marwick & boyd, 2010, p. 9). For instance, digital trace data collected from a particular social media or online platform may not represent a single context but rather multiple ones that are more, or less, loosely connected to one another. For example, the Reddit platform hosts multiple online communities that constitute very different contexts in which participants and participation patterns may be at odds with each other (Kitchens et al., 2020). Collecting digital trace data therefore requires an awareness of contexts that can be weakened through their collection.

3.1.2 Limited Situational Nuance

Related to the previous point, researchers who collect digital trace data may struggle with going beyond a limited understanding of a context. Interpreting digital traces can be challenging because of the absence of situational details (Marres & Gerlitz, 2016). Thus, researchers need to be cautious about inferring context from digital trace data, as these data may lack the critical elements of a situation (Pousti et al., 2021). Digital trace data can omit details, complicating their interpretation and analysis (Rava, 2022). Without questioning the context and looking for meaning, researchers risk oversimplifying complex social processes or misunderstanding what is happening (Bjerre-Nielsen & Glavind, 2022; Comunello et al., 2022).

3.1.3 Opacity of Algorithms Associated With Collected and Analyzed Digital Trace Data

Access to digital trace data is shaped by algorithms that filter and prioritize what users and researchers see (Gillespie et al., 2014). The invisible nature of such algorithmic interventions can make it hard to discern the actual context behind collected digital trace data (Ananny & Crawford, 2018). Algorithms can thus create artificial contexts that researchers may mistakenly consider to be real. This could thus lead them to theorize from a context that does not actually exist. Also, algorithms can produce digital trace data that only have a facade of objectivity (Leonardi & Treem, 2020). Further, algorithms can change, at times overnight and without researchers' awareness. This can lead to changes in researchers' access to and analysis of digital trace data. Therefore, the context researchers are investigating may not remain available to other researchers over time.

3.1.4 Ethics and Privacy Issues

Digital trace data are often publicly available, but collecting and analyzing them can create ethical issues related, for example, to people's privacy (Vaast & Walsham, 2013; Zimmer, 2020). Researchers face challenges in balancing the need for deep contextualization with respect for users' privacy, especially when traces may reveal personal or sensitive information. Since users may not expect their data to be used for research (Fiesler & Proferes, 2018), ethical considerations are crucial when contextualizing digital trace data, even when the data are, technically, publicly available.

3.2 Opportunities

However, digital trace data also offer opportunities for researchers, such as access to vast, often unobtrusive, and temporally marked records of behaviors and interactions. Adequately considered and analyzed, these data can help researchers theorize from contexts.

3.2.1 Access to Naturalistic Behavioral Data

Digital trace data offer researchers the ability to observe naturally occurring behaviors in real time (Lazer et al., 2009). Collecting digital trace data can give researchers access to behaviors and interactions as they occur organically (Eagle & Pentland, 2006). Digital trace data constitute "digital footprints" that researchers can follow (Golder & Macy, 2014) to gain an understanding of a context and to theorize from it.

3.2.2 Granularity of Observations and the Ability to Connect Micro and Macro Patterns

Digital trace data are often granular. Researchers can capture actions and interactions, as well as their time stamps and geolocations (Evans & Foster, 2011; Zhang et al., 2020). Researchers can then analyze digital trace data to identify and make sense of patterns at the microlevel (e.g., at the individual or small group level) (Mettler, 2024) as well as the macrolevel (e.g., social movements, shifts in public opinions) (Syed & Silva, 2023; Vaast et al., 2017). Digital trace data can be helpful for researchers seeking to elucidate multilevel theories and develop explanations for how microlevel actions scale at the macrolevel.

3.2.3 Real-Time and Dynamic Data Collection

Digital trace data allow researchers to follow shifts in behavior and content over time. They offer insights into the evolution of user behavior, changes in attitudes, and the dynamics of social settings that would be much more challenging for researchers to access otherwise (Hartl et al., 2023). This temporal richness can help researchers develop theories that account for changes in context, providing a deeper understanding of how behaviors and interactions unfold temporally (Kishore et al., 2024).

3.2.4 New Contexts, New Phenomena, and New Actors

Digital trace data can also be essential sources for researchers to access, understand, and then theorize from new, digitally enabled contexts and their dynamics (Johns, 2018; Vaast & Pinsonneault, 2022). With the advent of new technologies, social lives have changed, and new phenomena have arisen and become theorized— for instance, connective action (Bennett & Segerberg, 2012) and IT identity (Carter & Grover, 2015). Digital trace data can help researchers gain an understanding of these new phenomena and situate them in their contexts of occurrence (Armstrong et al., 2023). Without collecting and analyzing digital trace data, it would be challenging for researchers to develop an understanding of the contexts and to theorize from them (Shaikh & Vaast, 2023).

3.2.5 Access to Sensitive Topics

Digital trace data allow for the study of sensitive topics (e.g., health, mental well-being, political beliefs) and behavior in an organic manner (Andalibi et al., 2017; Murthy, 2024). Analyzing such topics through digital trace data can limit the risk of observer bias and the observer effect. The observer effect happens when something changes because of the presence of an observer, while observer bias is triggered when what is observed is affected by who is doing the observing. Accessing and analyzing digital trace data can also enable researchers to elaborate theories that account for sensitive topics with limited harm to participants (Andalibi et al., 2018). Access to digital trace data can be relatively safe for participants because the data are de facto generated and often publicly available.

To leverage this opportunity, researchers' sensitivity and carefulness are still required (see the challenges related to privacy and ethics if there is an expectation from users that the data is not actually public) (Fiesler & Proferes, 2018). Researchers thus still need to be aware of and reflect on the sensitivity of the issues they are investigating. They can also anonymize digital trace data when deemed necessary (Vaast, 2023).

Given these challenges and opportunities, this study proposes a twofold approach to theorize from contexts with digital trace data: first probing and then elucidating contexts that are associated with digital trace data.

4 Probing Contexts With Digital Trace Data

Probing contexts involves questioning the contexts under investigation and the ways in which digital trace data relate to them. As they probe contexts, researchers engage in surfacing the omnibus context and identifying, scanning, and connecting the discrete contexts (see Figure 1).

4.1 Surfacing the Omnibus Context

Johns (2006) distinguished between omnibus and discrete contexts. The omnibus context is broad and provides a general backdrop for discrete contexts to appear and phenomena to unfold. Discrete contexts are more specific and constitute the more immediate settings in which investigations are situated. The omnibus context may not always be at the forefront of investigations, but it is important to recognize because it gives particular meaning to events and actions (Davison & Martinsons, 2016). Thus, while the focus of a research project is on

discrete contexts, researchers need to acknowledge the omnibus context as well because it can drastically affect the digital trace data and findings. For instance, social media uprisings need to be situated within particular omnibus contexts (Faxon et al., 2023). The "MeToo" phenomenon was global, but its expressions, processes, and outcomes varied widely in different countries (Langer et al., 2020; Lopez et al., 2019).

Surfacing the omnibus context can help researchers situate digital trace data in the broader social, political, and economic environment of their production. Surfacing the omnibus context is important because digital trace data are produced and collected within particular environments that affect what these data say and how researchers interpret them. It is essential to avoid the "context collapse" (Marwick & boyd, 2010) of research relying on digital trace data. Surfacing the omnibus context can help researchers avoid misinterpretation regarding the digital trace data, the discrete contexts, and the phenomenon (Rava, 2022). It can also help them avoid erroneous conclusions, overstatements, or wrong attributions of causality in their theorizing.

4.2 Identifying Discrete Contexts

As they identify discrete contexts, researchers recognize what discrete contexts are directly related to the digital trace data. Identifying discrete contexts can help researchers situate their investigations and define the immediate settings from which the digital trace data emerge. Identifying involves clarifying how the digital trace data they collect come from particular contexts from which they will theorize.

In identifying discrete contexts, researchers define the boundaries of the contexts and determine which digital trace data are needed for the analysis and which are not, which is important, given how plentiful and highly granular digital trace data can be (Lindberg, 2020). Further, identifying discrete contexts may require researchers to engage in iterative sampling (Berente et al., 2019), i.e., to revisit their initial sampling as they question the discrete contexts of their research and determine which digital trace data they need. This can lead researchers to collect new data or to strategically set aside some of the data that are not relevant to the identified discrete contexts (e.g., Vaast et al., 2017)

Moreover, as they identify these contexts, researchers question whether their investigations relate to a single context or to more than one context. If they identify more than one context, they then ask what the different contexts under investigation are and how digital trace data are produced in these different contexts.

Also, along with the identification of contexts, researchers need an awareness of how the digital trace data are produced within the discrete contexts. This matters because no algorithm is neutral. Researchers can explore how the algorithms they rely on to collect digital trace data participate in shaping or distorting the contexts. In addition to identifying the contexts, researchers also need to assess whether the digital trace data they are collecting are related to sensitive topics and phenomena. This will also lead them to consider whether the participants and/or actions presented in the digital trace data and observed in the contexts require privacy protection.

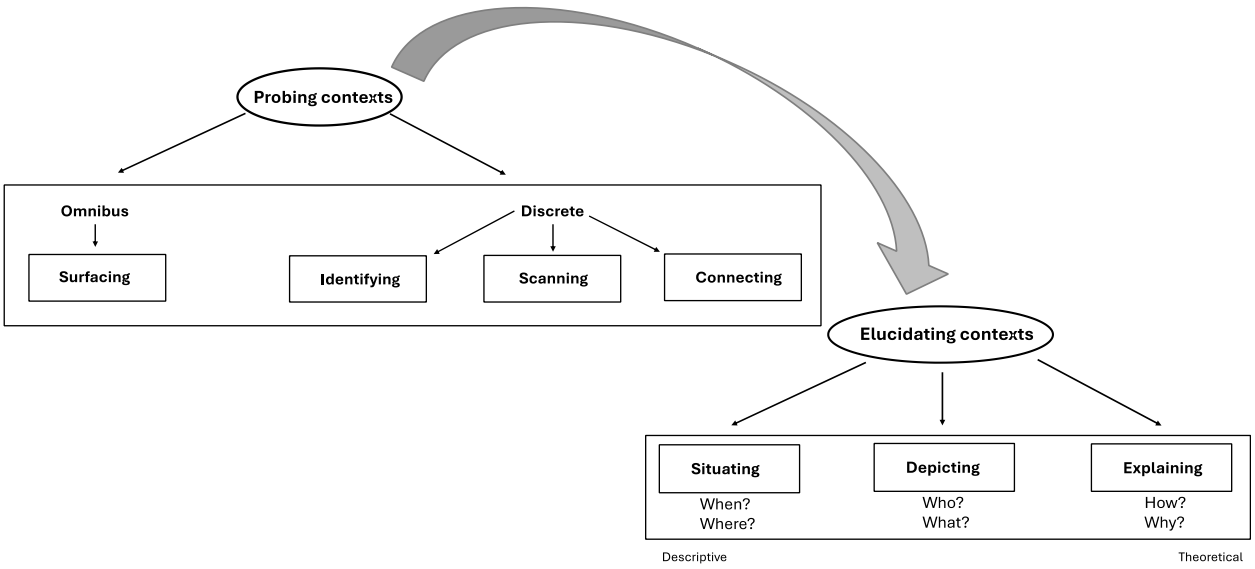

Figure 1. Probing and Elucidating Contexts With Digital Trace Data

4.3 Scanning Discrete Contexts

Scanning involves an initial exploration before making more targeted analytical moves. Scanning enables researchers to get a sense of the discrete contexts associated with the digital trace data. As they scan the discrete contexts, researchers need to ask whether the contexts are familiar to scholarship. Depending on the existing scholarship in various disciplines, different discrete contexts will be more, or less, familiar. For instance, organizational contexts such as companies or departments in organizations tend to be well-known to researchers in IS and management. Other contexts may be less so. Researchers should also ask whether the participants in these contexts are familiar. Bots, for instance, may act as participants in digitally enabled contexts. They can interact with other participants, and their actions can contribute to shaping and reshaping the contexts (Salge et al., 2022).

If the discrete contexts and/or participants are not well-known and understood, researchers may need to collect additional data to complement the digital trace data (e.g., semi-directed interviews). Additional data can help researchers develop a thicker understanding of new or underexamined contexts observed through digital trace data. Further, as they confront contexts that are relatively unknown, researchers can engage in iterative sampling of the digital trace data, whereby they can expand or refine the digital trace data they collect to gain a better understanding of the contexts (Berente et al., 2019).

Moreover, if the discrete contexts are new or poorly understood, researchers may face additional challenges in their theorizing. In such cases, researchers will need to delve deeper into the specifics of the discrete contexts to make sense of them and the phenomena they host. This can make generalizations in the theorizing more delicate, but it can also enable researchers to theorize about new phenomena. More generally, as they scan discrete contexts, researchers may also need to assess how new the contexts, phenomena, and actors that they are accessing through digital trace data are.

Scanning discrete contexts also involves determining what degree of granularity is needed for the analysis. Digital trace data can enable researchers to get very detailed in their analysis. This is especially the case for "exhaust data" (Kitchin, 2014), i.e., byproducts of digital activity such as time stamps, geolocations, and user interaction data that are automatically produced in sociotechnical systems. When digital trace data are highly granular, researchers may need to ask themselves whether they should retain all the data or sample them. The more granular the data, the more detailed the analysis and insights may become. However, this can make the overall understanding of the contexts and the theorizing exercise more delicate (Choi, 2020).

4.4 Connecting Discrete Contexts

If several discrete contexts are being investigated, researchers need to determine how these contexts are connected to one another and how the digital trace data relate to them. The discrete contexts observed through digital trace data may, for instance, be different but comparable on some important analytical dimensions. This connection among discrete contexts would then be reminiscent of the case study strategy of multiple comparative cases in traditional scholarship (Yin, 2009).

Another possibility is that the discrete contexts may overlap. The overlap may be temporal if the digital trace data refer to events and happenings within the same time frame. If certain participants are present in multiple discrete contexts, the overlap may also be one of actors and their actions (e.g., Vaast & Pinsonneault, 2022). This is particularly relevant for "platform ecologies" (Tufekci, 2017). Since participants can operate across multiple platforms, digital trace data that originate from one platform can also migrate to other platforms. Connecting discrete contexts is thus important because it highlights the need for researchers to analytically differentiate discrete contexts but also shows the importance of being open to relationships and porosity among these contexts (Zannettou et al., 2019).

Moreover, as they connect discrete contexts and examine how the digital trace data relate to them, researchers should take Miranda et al.'s (2022a) sampling stopping rules into consideration. These stopping rules can help researchers evaluate whether they need to collect additional digital trace data to compare and contrast the different discrete contexts they are investigating.

4.5 Illustrations

To illustrate the different aspects related to probing contexts with digital trace data, this study draws from three papers that relied on digital trace data, examined widely distinct contexts, and engaged in theorizing: Miranda et al. (2022b), Lindberg et al. (2024), and Pentland et al. (2021). Table 1 provides some summary elements of these papers. Table 2 highlights how these papers probe contexts.

Surfacing the omnibus context can help researchers develop an understanding of the overall background of their study (e.g., open source software development in Lindberg et al., 2024). Notably, each of the illustrative papers focused on digital trace data collected from English-speaking sources. Questioning these sources can help researchers uncover particular institutional contexts that can help them make sense of the collected digital trace data. Pentland et al.'s (2021) work is situated in the omnibus context of the US healthcare system, which has become digitized with the growing adoption of electronic medical records (EMRs). The healthcare sector's push for efficiency and standardization participates in the overall environment, with process dynamics unfolding within individual clinics.

Table 1. Summary Elements of the Illustrative Papers

	Miranda et al. (2022b)	Lindberg et al. (2024)	Pentland et al. (2021)
Research question	How can an innovation community's framing discourse be both diverse and coherent?	How do OSS communities use discourse to shape the novelty and complexity of the software they develop?	How does process change when participants are unaware of the change?
Foundations	• Discourse and framing theories • Innovation communities • Orders of worth	• Novelty and complexity in open source software development • Modulation (filtering, mixing)	• Process theorizing • Process modeling • Processes as directed graphs
Key findings / theorizing	Diversity and coherence coevolve as the diversity induced by mediated fields increases framing redundancies, synthesizing frames into a coherent community understanding of the innovation.	Modulation results in alternative OSS community approaches to shaping software novelty and complexity. This process reflects and is reflected in the resulting software artifact.	Identifying and measuring process changes uncovers shifts that are invisible to participants in the process.

Table 2. Probing Contexts With Digital Trace Data in the Illustrative Papers

Probing	Miranda et al. (2022b)	Lindberg et al. (2024)	Pentland et al. (2021)
Surfacing omnibus context: Examining the overall background of which digital trace data are part	English-speaking innovation community	Open source software development	US healthcare system
Identifying discrete contexts: Defining the discrete contexts, elaborating on their boundaries, and questioning how digital trace data are related to them	***Seven discursive fields of blockchain:*** Government, news media, science, idea evangelists, corporations, projects, social influencers ***Digital trace data:*** Articles, social media posts and profiles, websites, official documents	***Two OSS communities:*** Django, Rails ***Digital trace data:*** email newsletters, email threads, codebases for both communities	***Four dermatology clinics:*** From the University of Rochester Medical Center ***Digital trace data:*** Electronic medical records, time-stamped data from 57,000 patient visits
Scanning discrete contexts: Engaging in an initial exploration of the discrete contexts, determining the degree of granularity needed for the analysis	Six years (2012-2017)	Six years (2011-2016) ***Additional data:*** 25 interviews (12 for Django, 13 for Rails)	Two years (2016-2017) ***Additional data:*** Informal interviews with the clinical staff of the clinics
Connecting discrete contexts: If the digital trace data are related to more than one discrete context, questioning the relationships among these contexts	Distinct but related discursive fields on the blockchain innovation Combined contexts into: mediated fields, hybrid fields, and enactment fields	Two distinct contexts OSS communities selected to be similar in size and maturity, to be compared to each other	Four clinics from the same institution Data from the four clinics were contrasted and then combined

Identifying discrete contexts allows researchers to define and elaborate on the boundaries of the discrete contexts from which the digital trace data emerge. Miranda et al. (2022b) identified seven discursive fields (e.g., governments, news media, corporations) as distinct spaces where blockchain discourse is generated, each with its own framing logics. Lindberg et al. (2024) focused on two open source communities as discrete contexts where community members engage in technical debates and decision-making through mailing lists and developer forums.

Scanning the discrete contexts involves engaging in an initial exploration of the discrete contexts, which determines the degree of granularity needed for the analysis. Lindberg et al. (2024) explored discourse in the two communities and examined how it relates to novelty and complexity in development. Pentland et al. (2021), scanning the four clinics, maintained a very high degree of granularity in the observations, retaining the 57,000 data points to understand precisely the process changes over time.

These studies engaged in connecting discrete contexts to explore interdependencies and patterns across settings. Miranda et al. (2022b) uncovered how frames travel and mutate across discursive fields, illustrating how mediated fields act as bridges that enable frame migration and synthesis. Lindberg et al. (2024) contrasted two communities to highlight how distinct discursive modulation practices lead to different outcomes in software design. Pentland et al. (2021) connected four clinics by comparing process dynamics across them, revealing systemic changes that cut across individual clinics.

4.6 Transitioning Between Probing and Elucidating Contexts With Digital Trace Data

The transition between probing and elucidating contexts with digital trace data begins once researchers have developed an initial understanding of the contexts and are ready to start unpacking them for theorizing. Probing opens the field of inquiry and helps researchers sharpen their research focus, which is then deepened through elucidation. Miranda et al.'s (2022a) stopping rules, which indicate when researchers can move from probing to elucidating contexts, lie in the saturation of the initial examination of contexts. Saturation emerges from surfacing omnibus contexts and identifying, scanning, and connecting discrete contexts, which allows researchers to sufficiently explore their contexts to clear the way for elucidating.

For example, Miranda et al. (2022b) initially identified and contrasted seven discursive fields as distinct but connected discrete contexts shaping the blockchain discourse. Having probed their contexts, they then elucidated them by analyzing the framing diversity and coherence across contexts.

Of note, although Figure 1 presents a single arrow from probing to elucidating, the process between the two may not be simply linear. Certainly, probing helps researchers set the stage, while elucidating leads to theorizing. However, in practice, the process of probing and elucidating contexts may turn out to be iterative, as the different steps associated with elucidating contexts may lead researchers to return to some aspects of probing.

5 Elucidating Contexts With Digital Trace Data

Elucidating contexts observed through digital trace data involves situating, depicting, and explaining contexts. To do so, researchers must tackle key contextual questions of *where* and *when*, *what* and *who*, and *how* and *why*. Table 3 provides examples of some of the inquiries associated with these questions and the following paragraphs expand on this. These questions range from descriptive to explanatory and can help researchers progressively refine or elucidate new theory as they analyze their digital trace data. Notably, the following paragraphs discuss these issues sequentially. However, in research projects, researchers may tackle them in a closely related manner. See Figure 1 and Table 3 for a summary of how to elucidate contexts with digital trace data.

5.1 Situating Contexts: Where and When

Situating contexts involves answering *where* and *when* questions to spatiotemporally position the investigated contexts and the associated digital trace data. Collecting and analyzing digital trace data provide researchers with some clues to answer the *where* question of the contexts. Some digital trace data provide researchers with metadata such as geolocation (Zhang et al., 2020), allowing researchers to place the data within particular geographical and cultural settings. More abstractly, digital trace data can enable researchers to see how participants behave in ways that are more, or less, similar and, from this, infer a common context (Vaast & Pinsonneault, 2022). Digital trace data can provide signals for researchers to identify and make sense of the sociotechnical boundaries of the investigated contexts. Networks of connections among participants can also be indicative of the boundaries of shared contexts. Defining the *where* of the contexts is important for understanding what the theorical elaboration will be about. It also sets the stage for researchers to delimit the generalizability of their midrange theorizing.

As they collect and analyze digital trace data, researchers also temporally situate their investigations and answer *when*-related questions. Features of digital trace data such as time stamps, provide the exact timing of interactions, posts, edits, etc., and thus help researchers place observations in calendar time. On a diachronic level, through access to activity logs and digital archives, researchers can analyze digital trace data to identify temporal patterns and their relative frequencies, which would be difficult to observe otherwise (Pentland et al., 2021). Researchers can then identify and interpret evolutionary trends, shifts, bursts of activity, and lulls over time. For instance, from their collection of digital trace data, Tarafdar and Kajal Ray (2021) examined patterns of social media during an eight-day social protest cycle. Researchers can also rely on digital trace data as indicators of "flows" (Mousavi Baygi et al., 2021) and ongoing sociotechnical transformations.

Table 3. Examples of Questions to Elucidate Contexts Observed Through Digital Trace Data

		Examples of digital trace data and their analysis	Examples of questions
Situating contexts	**Where**	• Geolocation • Platform features	• Where does action take place, and how does technology shape the situation? • Where are the contexts situated, and what explains their formation?
	When	• Time stamps • Digital archives	• When do changes in actions or the environment occur? • When does context shift or evolve, and what triggers these changes?
Depicting contexts	**What**	• Behavioral data (e.g., likes, shares, comments) • Activity logs	• What is happening in the contexts? • What actions participate in the construction or transformation of the contexts?
	Who	• User IDs • Bot detection	• Who are the main participants in the contexts? • Who is participating in different ways in the contexts?
Explaining contexts	**How**	• Network analysis • Process modeling • Correlation, causal inferences • Sentiment analysis	• How do participants' actions shape and reshape the contexts? • How does the context emerge and evolve?
	Why		• Why do agents engage with the contexts in particular ways? • Why do observations from the digital trace data change over time?

Asking *when* questions can therefore help researchers temporally define contexts and differentiate between synchronic and diachronic investigations. This is especially important because contexts change (Grisold et al., 2024b); they can be stable for a time but then become fluid and evolve (McLaren & Durepos, 2021). This is particularly the case for contexts observed through digital trace data. Drechsler et al. (2022), for instance, relied on digital trace data to examine the evolution over time of a digital infrastructure. Answering *when* questions can thus help researchers develop an understanding of the temporal factors and transformations associated with the investigated contexts that shape behavior and a phenomenon of interest. To help answer *when* questions in a longitudinal manner, researchers can seek to explain transformations over time through techniques such as temporal bracketing (Hartl et al., 2023) and process modeling (Pentland et al., 2021).

5.2 Depicting Contexts: What and Who

Depicting contexts involves developing a deeper understanding of contexts and digital trace data by unpacking what is happening and who is participating.

Addressing *what* questions can help researchers define and understand the phenomenon they are investigating (Monteiro et al., 2022). Digital trace data offer researchers numerical information, text, images, videos, or other data that reveal *what* is being shared, done, or discussed in the contexts observed. As they ask *what* questions, researchers narrow in on the actions and happenings in the contexts they investigate. Since the data they collect provide traces of these actions and happenings, researchers can delve into what is happening with varying degrees of granularity.

Regarding *who* questions, collecting and analyzing digital trace data allows researchers to question which agents (human and/or technological) are involved in creating and engaging with digital trace data. Pseudonyms, usernames, and IDs can provide answers to *who* questions. Moreover, network data and interaction logs (e.g., on social media, measures of engagement such as likes, shares, and comments) can reflect engagement relations among agents. While staying mindful of ethical issues that may arise (Vaast & Urquhart, 2017; Vaast & Walsham, 2013), researchers can question who the participants are in the contexts they investigate, as well as what roles they play and how they interact.

Depicting contexts by answering *what* and *who* questions can help researchers clarify the novelty of the phenomenon and the different and new character of the participants in the contexts they examine. It refines what theorizing from contexts is about—e.g., bot behavior on social media (Nyman et al., 2024) or algorithmic management and worker negotiation (Salge et al., 2022).

5.3 Explaining Contexts: How and Why

As they continue to deepen their understanding of contexts, researchers engage in further analysis of digital trace data to answer questions related to how and why actions and events unfold the way they do.

Answering *how* questions can help researchers elucidate how actions happen in contexts observed through digital trace data. To answer *how* questions, researchers can build on their digital trace data to develop process tracing or modeling on the event-sequence or network analysis to reveal how actions unfold over time within and across these contexts. For instance, by modeling temporal patterns of commits, discussions, or decisions, researchers can trace mechanisms linking actions to outcomes (Kyriakou et al., 2017). Answering *how* questions can help researchers reveal emergent patterns happening in the contexts from their analysis of digital trace data.

Answering *why* questions is at the core of theorizing for researchers. Raw digital trace data are unlikely to help researchers address these important questions, but further analysis of digital trace data can help researchers work towards addressing *why* questions. As they ask *why,* researchers examine why observations and actions align with or deviate from what they expected and question the roles of the contexts in this alignment and deviation. Analyzing correlations and running causal inference techniques on large datasets of digital trace data can help researchers understand events and behaviors. Sentiment analysis can also give researchers glimpses of the emotional state of participants in context as events happen and contexts evolve.

Addressing *how* and *why* questions is essential for researchers as they work at theorizing from contexts. This enables them to contextualize their analysis of digital trace data while building theory that illuminates the contexts and phenomena under investigation.

5.4 Illustrations

Table 4 summarizes the ways in which the three illustrative papers address the elucidating context questions. The following paragraphs elaborate on some notable aspects.

The illustrative papers engaged in situating the contexts. Miranda et al. (2022b) positioned blockchain discourse within a global sociotechnical ecosystem by examining seven distinct discursive fields, such as governments and corporations, across the six-year period (2011-2017) in which the blockchain hype was rapidly growing. Pentland et al. (2021) focused on four dermatology clinics using EMR logs to observe process dynamics over a two-year timespan, illustrating how time-stamped trace data reveal hidden process changes in workflows.

Regarding depiction, the papers highlight how the digital trace data are related to different occurrences and the related participants. For instance, Lindberg et al. (2024) depicted discursive practices in each community, unpacking what types of debates and decisions occurred and who drove them. Miranda et al. (2022b) explored what frames regarding blockchain emerged from various fields, identifying government officials, media actors, and corporate strategists as key participants shaping the meaning of blockchain over time.

The studies also engaged in explaining how and why events happened and evolved the way they did in the focal contexts. Pentland et al. (2021) explained how workflows evolved by tracing changes in process structures via directed graphs, showing why seemingly stable clinical routines experienced unseen shifts over time. Lindberg et al. (2024) explained how two communities modulated discourse via filtering and mixing practices to balance software novelty and complexity and showed why the two communities took distinct paths in shaping their technical artifacts.

5.5 Elucidating Contexts and Theorizing From Them

Elucidating contexts constitutes an important methodological stance that allows researchers to theorize from contexts. It guides researchers in their theorizing in that the elucidating questions inform the theoretical insights that researchers can derive. Specifically, answers to *when* and *where* questions situate the insights. Answers to *what* and *who* questions clarify the participants and actions. Answers to *how* and *why* questions illuminate emergent patterns and their explanations. Together, these answers contribute to theory development.

For example, Pentland et al. (2021) identified key actors in the dermatology appointment process and unpacked how and why the process evolved over a two-year period. As they elucidated how the process changed over time, they theorized how small workflow changes that are often invisible can have larger effects in healthcare routines and delivery.

Elucidating thus enables theorizing from contexts by helping researchers develop nuanced situated insights that enable them to theorize in a way that is both grounded in the contexts and generalizable through abstraction.

Table 4. Elucidating Contexts With Digital Trace Data in the Illustrative Papers

Elucidating contexts	Miranda et al. (2022b)	Lindberg et al. (2024)	Pentland et al. (2021)
Situating: Developing an understanding of the circumstances of the contexts by answering when and where questions	***Where:*** English-language online discourse ***When:*** Discourse related to blockchain from when it started appearing (2012) to the peak of the hype cycle (2017).	***Where:*** Open-source software development in two communities ***When:*** 2011-2016	***Where:*** Dermatology appointments in four dermatology clinics ***When:*** Time stamps of appointment activities, 2016-2017
Depicting: Digging into what is happening in the contexts and who participates in these happenings	***What:*** Discourse on blockchain ***Who:*** Participants in seven discursive fields, combined into three categories	***What:*** Discursive modulation practices ***Who:*** Participants in the two communities	***What:*** Events associated with patient visits ***Who:*** Patients, medical staff, clerical staff
Explaining: Building elements of answers to understand how and why events happen and evolve in the contexts	***How:*** Discourse on blockchain evolves over time, displaying both diversity and coherence. ***Why:*** Framing diversity and coherence come from different but connect discursive fields.	***How:*** Discursive modulation practices differ in the two communities. ***Why:*** Different modulation practices are associated with distinct measures of novelty and complexity.	***How:*** The appointment process changes in particular temporally defined patterns. ***Why:*** The appointment process changes over time due to distinct factors even if the staff is unaware of it.

6 Discussion and Implications

6.1 Discussion

This paper provides a framework for researchers to theorize from contexts observed through digital trace data. Digital trace data enable researchers to gain in-depth access to events and happenings in multiple contexts. Researchers may face challenges as they balance the need for contextualizing and theorizing when engaging with digital trace data. This study suggests that first probing and then elucidating contexts can help researchers deal with these challenges. This discussion builds on two considerations that emerged from applying this framework. The first is related to the importance of surfacing the omnibus context to identify boundary conditions and opportunities for further research. The second is related to delimiting and connecting one or more discrete contexts when analyzing digital trace data.

First, surfacing the omnibus context can help researchers both position their inquiry and define the limits of their theorizing. For instance, Miranda et al. (2022b) surfaced an omnibus context of the English-speaking innovation discourse, implicitly shaped by Western institutional logics and discourse traditions. While this provides a rich understanding of how blockchain is framed within this large ecosystem, it also suggests a research opportunity to examine non-Western or non-English discourse. Surfacing the omnibus context makes visible the often-implicit assumptions within the sociotechnical environments from which digital trace data arise (Davison & Martinsons, 2016). As such, this step is essential not only for refining the scope of midrange theorizing but also for identifying paths to extend investigations to alternative omnibus contexts, which may exhibit distinct framing logics, power asymmetries, and/or discursive norms.

Second, discrete contexts are rarely isolated when traced through digital environments. Lindberg et al. (2024) explicitly focused on two open source communities as separate contexts. However, their comparison highlights how similar governance structures can yield divergent discursive modulation practices. Pentland et al. (2021) similarly studied four clinics as distinct yet connected contexts within the same institutional system. The ability to detect relationships or overlaps among discrete contexts is key to explaining how patterns emerge and evolve across contexts. In digital research, where platform ecologies often create porous boundaries (Tufekci, 2017; Zannettou et al., 2019), identifying and connecting discrete contexts can help researchers avoid the pitfalls of treating each setting as hermetically sealed. Yet, such cross-contextual dynamics can make the task of delimiting context boundaries analytically challenging when working with digital trace data.

These considerations also underscore the risks inherent in theorizing from digital trace data. Digital environments evolve rapidly and often unpredictably (Dourish, 2004), making it difficult to build theories that remain relevant as sociotechnical conditions shift. A theory grounded in specific platform configurations, for example, may lose explanatory power when those configurations change due to new algorithms, governance practices, or user behaviors. Moreover, scholars may risk overstating the novelty of digital contexts and phenomena. Without carefully probing and elucidating contexts, there is a danger of emphasizing apparent novelty while overlooking relevant existing theories that could enrich the explanation.

In this light, the value of probing (surfacing, identifying, scanning, connecting) and elucidating (situating, depicting, explaining) contexts is not only methodological but also epistemological. These activities can help researchers develop theories that are sensitive to the sociotechnical specificities of digital environments while also abstracting from these idiosyncrasies. Additionally, they can provide a safeguard against misinterpretation, context collapse, and overgeneralization. In sum, theorizing from contexts with digital trace data requires researchers to reflect on the appropriate balance between depth and generalizability, between embracing the complexity of contexts and distilling theoretical insights that transcend them.

6.2 Implications

Digital trace data have become useful inputs for many research projects, but researchers face dilemmas when trying to theorize from contexts observed through digital trace data. This paper provides a framework for researchers to define and elaborate on how to theorize from contexts observed through digital trace data.

Much scholarship on digital trace data has focused on these data and on their limitations and the opportunities they offer researchers (Freelon, 2014; Howison et al., 2011). Against this backdrop, bringing the context back into research relying on digital trace data is useful because it can help researchers think about what collecting and analyzing these data may reveal and how it can help them theorize. This is especially important because part of the appeal of digital trace data relies on their relative ease of access. Compared to, say, data collected through interviews or participant observations or, even, non-digital archives, digital trace data are convenient to collect. However, such convenience does not guarantee that digital trace data represent interesting contexts that researchers can use to theorize. Scholars have offered methodological guidance on dealing with digital trace data (e.g., Lindberg, 2020; Østerlund et al., 2020), and the methodological innovation of computationally intensive theory construction (Berente et al., 2019; Miranda et al., 2022a) has provided great insights into how to build theory using digital trace data. Such recommendations are invaluable for researchers seeking to understand how to handle and analyze these data. However, these recommendations have so far provided limited insights into how to deal with the contexts observed through collected digital trace data. This study thus complements and adds to the existing insights by shedding light on the importance of theorizing from contexts with digital trace data.

Bringing context to the forefront for research relying on digital trace data is also important because the process of research with digital trace data is itself affected by the investigated contexts. Patterns with theoretical implications (Miranda et al., 2022a), for instance, require a deep understanding of the contexts to which they apply. Without questioning and elaborating on contexts, the theoretical implications that researchers can propose are weakened. Furthermore, considering context also adds to scholarship focusing on phenomena and their problematization (Monteiro et al., 2022). Indeed, phenomena cannot be decontextualized. To deepen the understanding of a focal phenomenon, it is essential for researchers to consider context—especially when relying on digital trace data that can seem, at first glance, to lack context. Researchers thus need to define the *where/when*, *what/who*, and *how/why* of their contexts to assess how novel, different, and intriguing they are, and to elaborate theoretically on them.

Also, this study has implications for scholarship because it can help researchers take context seriously without losing sight of the goal of elaborating theory. This is important because of the apparent tension between digging deeply into contexts and being able to theorize in a way that accounts for but also transcends contexts (Dyer & Wilkins, 1991; Johns, 2006). This is also especially critical for research dealing with digital trace data, given the growing trend towards empiricism in computational social science (Lazer et al., 2009). The ambition of theorizing from contexts is indeed to develop new theory, thanks to research that relies on digital trace data rather than merely reports empirically on these data.

Furthermore, this study does not consider contexts as fixed once and for all. It highlights that contexts, especially those that are represented by digital trace data, are dynamic. They can change drastically and rapidly as technology and algorithms develop, platform policies shift, and participants' behaviors evolve. Theorizing contexts observed through digital trace data requires accounting for their fluidity (Dourish, 2004). This is exciting for researchers because it enables them to observe and develop an understanding of new dynamics, but it can also be overwhelming and can lead researchers to focus too much on ephemeral issues. The questions of *when* and *why*, in particular, can help researchers develop theory that captures both stability and flux in contexts observed through digital trace data.

7 Conclusion

In closing, because many researchers rely on—and are likely to continue relying on—digital trace data in their work, the importance of methodological guidelines that help researchers manage the opportunities and challenges of these data has grown. This paper provides guidelines on how to elaborate theory using digital trace data in a rigorous manner and adds to previous insights by emphasizing the importance of being sensitive to contexts when performing research involving digital trace data. Theorizing from contexts can help researchers contextualize their insights as they collect and analyze digital trace data while also allowing them to develop new theory.

References

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973-989.

Andalibi, N., Haimson, O. L., Choudhury, M. D., & Forte, A. (2018). Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. *ACM Transactions on Computer-Human Interaction, 25*(5), Article 28.

Andalibi, N., Ozturk, P., & Forte, A. (2017). Sensitive self-disclosures, responses, and social support on Instagram: The case of #Depression. *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing.*

Armstrong, A., Briggs, J., Moncur, W., Carey, D. P., Nicol, E., & Schafer, B. (2023). Everyday digital traces. *Big Data & Society, 10*(2). https://doi.org/10.1177/20539517231213827

Ash, J., Leszczynski, A., & Kitchin, R. (2024). *Researching digital life: Orientations, methods and practice*. SAGE.

Avital, M., Jensen, T. B., & Dyrby, S. (2023). The social fabric framework: steps to eliciting the social making of organisations in the digital age. *European Journal of Information Systems, 32*(2), 127-153.

Baiyere, A., Grover, V., Lyytinen, K. J., Woerner, S., & Gupta, A. (2023). Digital "x"—Charting a path for digital-themed research. *Information Systems Research, 34*(2), 463-486.

Bamberger, P. (2008). From the editors beyond contextualization: Using context theories to narrow the micro-macro gap in management research. *Academy of Management Journal, 51*(5), 839-846.

Baskerville, R. L., Myers, M. D., & Yoo, Y. (2020). Digital first: The ontological reversal and new challenges for IS research. *MIS Quarterly, 44*(2), 509-523.

Beer, D., & Burrows, R. (2007). Sociology and, of, and in Web 2.0: Some initial considerations. *Sociological Research Online, 12*(5), 67-79.

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, Communication & Society, 15*(5), 739-768.

Berente, N., Seidel, S., & Safadi, H. (2019). Research commentary—Data-driven computationally intensive theory development. *Information Systems Research, 30*(1), 50-64.

Bjerre-Nielsen, A., & Glavind, K. L. (2022). Ethnographic data in the age of big data: How to compare and combine. *Big Data & Society, 9*(1). https://doi.org/10.1177/20539517211069893

boyd, D., & Crawford, K. (2011). Six provocations for big data. *Proceedings of the Symposium on the Dynamics of the Internet and Society*.

Breiter, A., & Hepp, A. (2018). The complexity of datafication: putting digital traces in context. In A. Hepp, A. Breitner, & U. Hasekbrink (Eds.). *Communicative figurations* (pp. 387-405). Palgrave Macmillan.

Carter, M., & Grover, V. (2015). Me, my self, and I(T): Conceptualizing information technology identity and its implications. *MIS Quarterly, 39*(4), 931-958.

Choi, S. (2020). When digital trace data meet traditional communication theory: Theoretical/ methodological directions. *Social Science Computer Review, 38*(1), 91-107.

Comunello, F., Martire, F., & Sabetta, L. (2022). *What people leave behind: Marks, traces, footprints and their relevance to knowledge society*. Springer.

Da Cunha, J. V., & Orlikowski, W. J. (2008). Performing catharsis: The use of online discussion forums in organizational change. *Information and Organization, 18*, 132-156.

Davison, R. M., & Martinsons, M. G. (2016). Context is king! Considering particularism in research design and reporting. *Journal of Information Technology, 31*(3), 241-249.

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing, 8*(1), 19-30.

Drechsler, K., Grisold, T., Gau, M., & Seidel, S. (2022). Digital infrastructure evolution: A digital trace data study. *Proceedings of the International Conference of Information Systems.*

Dyer, W. G., & Wilkins, A. L. (1991). Better stories, not better constructs, to construct better theory: A rejoinder to Eisenhardt. *Academy of Management Review, 16*(3), 613-619.

Eagle, N., & Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing, 10*, 255-268.

Evans, J. A., & Foster, J. G. (2011). Metaknowledge. *Science, 331*(6018), 721-725.

Faxon, H. O., Kintzi, K., Tran, V., Wine, K. Z., & Htut, S. Y. (2023). Organic online politics: Farmers,

Facebook, and Myanmar's military coup. *Big Data & Society, 10*(1). https://doi.org/10.1177/20539517231168101

Fiesler, C., & Proferes, N. (2018). "Participant" perceptions of Twitter research ethics. *Social Media + Society, 4*(1). https://doi.org/10.1177/2056305118763366

Fisher, G., Mayer, K., & Morris, S. (2021). From the editors—Phenomenon-based theorizing. *Academy of Management Review, 46*(4), 631-639.

Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. *Journal of Broadcasting & Electronic Media, 58*(1), 59-75.

Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014). *Media technologies: Essays on communication, materiality, and society*.

Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology, 40*, 129-152.

Griffin, M. A. (2007). Specifying organizational contexts: systematic links between contexts and processes in organizational behavior. *Journal of Organizational Behavior, 28*(7), 859-863.

Grisold, T., Aa, H. v. d., Franzoi, S., Hartl, S., Mendling, J., & Brocke, J. V. (2024a). *Proceedings of the 6th International Conference on Process Mining.*

Grisold, T., Kremser, W., Mendling, J., Recker, J., Brocke, J. v., & Wurm, B. (2023). Keeping pace with the digital age: Envisioning information systems research as a platform. *Journal of Information Technology, 38*(1), 60-66.

Grisold, T., Kremser, W., Mendling, J., Recker, J., vom Brocke, J., & Wurm, B. (2024b). Generating impactful situated explanations through digital trace data. *Journal of Information Technology, 39*(1), 2-18.

Grover, V., Lindberg, A., Benbasat, I., & Lyytinen, K. (2020). The perils and promises of big data research in information systems. *Journal of the Association for Information Systems, 21*(2), 268-291.

Grover, V., & Lyytinen, K. (2015). New state of play in information systems research: The push to the edges. *MIS Quarterly, 39*(2), 271-296.

Grover, V., & Lyytinen, K. (2023). The pursuit of innovative theory in the digital age. *Journal of Information Technology, 38*(1), 45-59.

Hampton, K. N. (2017). Studying the digital: Directions and challenges for digital methods. *Annual Review of Sociology, 43,* 167-188.

Hartl, S., Franzoi, S., Grisold, T., & Vom Brocke, J. (2023). Explaining change with digital trace data: A Framework for temporal bracketing. *Proceedings of the Hawaii International Conference on Systems Sciences.*

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., & Dhillon, G. (2014). A framework and guidelines for context-specific theorizing in information systems research. *Information Systems Research, 25*(1), 111-136.

Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems, 12*(12), 767-797.

Jiang, D., Jiang, L. D., Jackie, J., Jr., Grover, V., & Sun, H. (2022). Everything old can be new again: Reinvigorating theory borrowing for the digital age. *MIS Quarterly, 46*(4), 1833-1850.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386-408.

Johns, G. (2018). Advances in the treatment of context in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior, 5*, 21-46.

Kishore, S., Sundaram, D., & Myers, M. D. (2024). A temporal dynamics framework and methodology for computationally intensive social media research. *Journal of Information Technology*, *40*(2), 140-163.

Kitchens, B., Johnson, S. L., & Gray, P. (2020). Understanding echo chambers and filter bubbles: The impact of social media on diversification and partisan shifts in news consumption. *MIS Quarterly, 44*(4), 1619-1649.

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

Kyriakou, H., Nickerson, J. V., & Sabnis, G. (2017). Knowledge reuse for customization. *MIS Quarterly, 41*(1), 315-332.

Langer, N., Khurana, S., & Vaast, E. (2020). Activating the sisterhood: A structural and temporal analysis of sustained connective action in #MeTooIndia. *Proceedings of the International Conference of Information Systems*.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science, 323*(5915), 721-723.

Leonardi, P. M., & Treem, J. W. (2020). Behavioral Visibility: A new paradigm for organization studies in the age of digitization, digitalization, and datafication. *Organization Studies, 41*(12), 1601-1625.

Lindberg, A. (2020). Developing theory through integrating human and machine pattern recognition. *Journal of the Association for Information Systems, 21*(1), 90-116.

Lindberg, A., Berente, N., Howison, J., & Lyytinen, K. (2024). Discursive modulation in open source software: How online communities shape novelty and complexity. *MIS Quarterly, 48*(4), 1395-1422.

Lopez, I., Quillivic, R., Evans, H., & Arriaga, R. I. (2019). Denouncing sexual violence: A cross-language and cross-cultural analysis of #MeToo and #BalanceTonPorc. *Proceedings of the 17th IFIP TC 13 International Conference*.

Marres, N., & Gerlitz, C. (2016). Interface methods: Renegotiating relations between digital social research, STS and sociology. *The Sociological Review, 64*(1), 21-46.

Marwick, A. E., & boyd, d. (2010). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society, 13*(1), 114-133.

McLaren, P. G., & Durepos, G. (2021). A call to practice context in management and organization studies. *Journal of Management Inquiry, 30*(1), 74-84.

Merton, R. K. (1957). *Social theory and social structure*. Free Press.

Mettler, T. (2024). The connected workplace: Characteristics and social consequences of work surveillance in the age of datification, sensorization, and artificial intelligence. *Journal of Information Technology, 39*(3), 547-567.

Miranda, S. M., Berente, N., Seidel, S., Safadi, H., & Burton-Jones, A. (2022a). Computationally intensive theory construction: A primer for authors and reviewers. *MIS Quarterly, 46*(2), i-xvii.

Miranda, S. M., Wang, D., & Tian, C. (2022b). Discursive fields and the diversity-coherence paradox: An ecological perspective on the blockchain community discourse. *MIS Quarterly, 46*(3), 1421-1452.

Monteiro, E., Constantinides, P., Scott, S., Shaikh, M., & Burton-Jones, A. (2022). Qualitative research methods in information systems: a call for phenomenon-focused problematization. *MIS Quarterly, 46*(2), i-xviii.

Mousavi Baygi, R., Introna, L. D., & Hultin, L. (2021). Everything flows: Studying continuous sociotechnological transformation in a fluid and dynamic digital world. *MIS Quarterly, 45*(1), 423-452.

Murthy, D. (2024). Sociology of Twitter/X: Trends, challenges, and future research directions. *Annual Review of Sociology, 50*, 169-190.

Nyman, S., Bødker, M., & Blegind Jensen, T. (2024). Reforming work patterns or negotiating workloads? Exploring alternative pathways for digital productivity assistants through a problematization lens. *Journal of Information Technology, 39*(3), 503-520.

Østerlund, C., Crowston, K., & Jackson, C. (2020). Building an apparatus: Refractive, reflective, and diffractive readings of trace data. *Journal of the Association for Information Systems, 21*(1), 1-22.

Pentland, B., Vaast, E., & Ryan Wolf, J. (2021). Theorizing process dynamics with directed graphs: A diachronic analysis of digital trace data. *MIS Quarterly, 45*(2), 967-984.

Pentland, B. T., Recker, J., Wolf, J. R., & Wyner, G. (2020). Bringing context inside process research with digital trace data. *Journal of the Association for Information Systems, 21*(5), 1214-1236.

Pike, J. C., Bateman, P. J., & Butler, B. S. (2018). Information from social networking sites: Context collapse and ambiguity in the hiring process. *Information Systems Journal, 28*(4), 729-758.

Pousti, H., Urquhart, C., & Linger, H. (2021). Researching the virtual: A framework for reflexivity in qualitative social media research. *Information Systems Journal, 31*(3), 356-383.

Rava, G. (2022). Traces and their (in)significance. In F. Comunello, F. Martire, & L. Sabetta (Eds.), *What people leave behind: Marks, traces, footprints and their relevance to knowledge society* (pp. 269-281). Springer.

Revina, A., Rizun, N., & Aksu, Ü. (2023). Towards a framework for context awareness based on textual process data: Case study insights. *Proceedings of the IDAMS, SoEA4EE, TEAR, EDOC Forum, Demonstrations Track and Doctoral Consortium*.

Rogers, R. (2013). *Digital methods*. MIT Press.

Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press.

Salge, C. A. d. L., Karahanna, E., & Thatcher, J. B. (2022). Algorithmic processes of social alertness and social transmission: How bots disseminate information on Twitter. *MIS Quarterly, 46*(1), 229-259.

Shaikh, M., & Vaast, E. (2023). Algorithmic interactions in open source work. *Information Systems Research*, *34*(2), 744-765.

Shevchuk, A., Strebkov, D., & Tyulyupo, A. (2021). Always on across time zones: Invisible schedules in the online gig economy. *New Technology, Work and Employment, 36*(1), 94-113.

Syed, R., & Silva, L. (2023). Social movement sustainability on social media: An analysis of the women's march movement on Twitter. *Journal of the Association for Information Systems, 24*(1), 249-293.

Tarafdar, M., & Kajal Ray, D. (2021). Role of social media in social protest cycles: A sociomaterial examination. *Information Systems Research, 32*(3), 1066-1090.

Tufekci, Z. (2017). *Twitter and tear gas: The power and fragilitty of networked protest*. Yale University Press.

Vaast, E. (2023). Strangers in the dark: Navigating opacity and transparency in open online career-related knowledge sharing. *Organization Studies, 44*(1), 29-52.

Vaast, E., & Pinsonneault, A. (2022). Dealing with the social media polycontextuality of work. *Information Systems Research, 33*(4), 1428-1451.

Vaast, E., Safadi, H., Lapointe, L., & Negoita, B. (2017). Social media affordances for connective action: An examination of microblogging use during the Gulf of Mexico oil spill. *MIS Quarterly, 41*(4), 1179-1205.

Vaast, E., & Urquhart, C. (2017). Building grounded theory with social media data. In R. Mir & S. Jain (Eds.), *Routledge companion to qualitative research in organization studies* (Ch. 27). Routledge.

Vaast, E., & Walsham, G. (2013). Grounded theorizing for electronically mediated social contexts. *European Journal of Information Systems, 22*(1), 9-25.

Venkatesh, V. (2025). Leveraging context: Re-thinking research processes to make "contributions to theory." *Information Systems Research*, *36*(3), 1871-1886.

Vial, G. (2019). Reflections on quality requirements for digital trace data in IS research. *Decision Support Systems, 126*, Article 113133.

Vicari, S., & Kirby, D. (2023). Digital platforms as socio-cultural artifacts: Developing digital methods for cultural research. *Information, Communication & Society, 26*(9), 1733-1755.

Welch, C., Paavilainen-Mäntymäki, E., Piekkari, R., & Plakoyiannaki, E. (2022). Reconciling theory and context: How the case study can set a new agenda for international business research. *Journal of International Business Studies, 53*(1), 4-26.

Whetten, D. A. (2009). An examination of the interface between context and theory applied to the study of Chinese organizations. *Management and Organization Review, 5*(1), 29-56.

Xu, H., & Zhang, N. (2022). From contextualizing to context theorizing: Assessing context effects in privacy research. *Management Science, 68*(10), 7383-7401.

Yin, R. K. (2009). *Case study research*. SAGE.

Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). Who let the trolls out? Towards understanding state-sponsored trolls. *Proceedings of the 10th ACM Conference on Web Science.*

Zhang, Y., Li, B., & Krishnan, R. (2020). Learning individual behavior using sensor data: The case of global positioning system traces and taxi drivers. *Information Systems Research, 31*(4), 1301-1321.

Zhang, Z., Lee, H., Yoo, Y., & Choi, Y. T. (2022). Theorizing routines with computational sequence analysis: A critical realism framework. *Journal of the Association for Information Systems, 23*(2), 589-630.

Zimmer, M. (2020). "But the data is already public": On the ethics of research in Facebook. In K. W Miller, & M. Taddeo (Eds.), *The ethics of information technologies* (pp. 229-241). Routledge.

About the Authors

Emmanuelle Vaast is a professor of information systems and the Desautels Chair in Digital Technology Management at the Desautels Faculty of Management of McGill University. She has served as a senior editor of *Information Systems Research* and *MIS Quarterly*. Her research examines how social practices emerge and change with the implementation and use of new technologies and how these new practices are associated with organizational and change dynamics. She is fascinated by the new practices and social and societal changes associated with digital technologies. Some of the themes she is especially interested in deal with the emergence of new organizational forms and the new dynamics associated with organizational and occupational identification and the future of work. Most of her research so far has been qualitative, relying on the analysis of case-based evidence. She has, however, become increasingly interested in the opportunities and challenges of theorizing with digital trace data and computations.