Information Research - Vol. 30 No. CoLIS (2025)

Being literate in data or practices: how paradata functions in the context of literacy

Isto Huvila, Michael Olsson, Olle Sköld, Jessica Kaiser, Lisa Andersson

DOI: https://doi.org/10.47989/ir30CoLIS52324

Abstract

Introduction. Datafication of information practices underlines the importance of better understanding of data literacy and documentation. Framing data literacy as a sociocultural information practice, we draw from its conceptual affinities with the notion paradata to advance the understanding what it takes to be(come) literate, and how paradata functions in the context of literacy.

Method. Theoretical discussion draws from an interviews with (N=) 33 professionals working with archaeological data.

Analysis. Interview transcripts were analysed using a method based on the constant comparative method and close reading.

Results. Four facets of being literate with archaeological data were identified: information making literacy, syntactic literacy, information taking literacy and interpretative literacy. Becoming literate is influenced by the parallel (in)adequacy of individual facets. Literacies' social being-in-making makes becoming literate hard. Finally, paradata potentially 1) contributes developing and enacting individual literacies, 2) catalysing and impeding convergence of literacies, and 3) enacting the contexts of literacies.

Conclusion. Paradata, as a concept referring to practice-related meta-information, provides means to operationalise what can be known and communicated about practices as a contextual constituent of literacy. Paradata helps in becoming and staying literate. Finally, literacy can be a question of competent practice with paradata rather than mastering the data itself.

Introduction

The rapid surge of datawork as a dominant modality of producing knowledge, and data as its fundamental ingredient, underlines the importance of better practical and conceptual understanding of data literacy. Much of the existing literature on data literacy follows the conventional framing of literacies as abilities and competencies in relation to *information*, or in the case of data literacy, to comparable informational entities such as data and metadata (Fernández-Pascual et al., 2024; Tuominen et al., 2005). In parallel, however, the literature on data practices is underlining the processual, relational and practice-based nature of data entangled in a sociomaterial world (e.g., Faniel et al., 2019; Kim and Yoon, 2017), very similarly to how information science research has increasingly shifted from the objectivist and individually oriented approaches to information (e.g., Cox, 2013; Olsson and Lloyd, 2017; Savolainen, 2008), recognising the comparably entangled nature of information and information practices. Considering this, we argue that research into data literacy would greatly benefit from being grounded in a conceptual framework that moves beyond skills-centred concepts of literacy prevalent in the literature and practices of the information professions, and examines data literacy as a complex, highly-contextualised social practice.

This paper takes as its starting point Lloyd's definition of information literacy as a sociocultural information practice which emerges through shared understandings, social prescriptions, beliefs and understandings specific to a site/community (Lloyd, 2010). In contrast to skills and competences-based approaches to (information) literacy, it follows the tenets of the new literacy studies (Gee, 2000) in how it emphasises the multiplicity of literacies as embodied and material practices rather than as a generic ability. A central consequence of adopting an information practice-based approach to data literacy is the need to expand our understanding of both data and the associated method and data descriptions that its users and re-users will need to make sense of data made by other data authors.

This approach to literacy both contrasts and parallels with the conceptual focus of the notion of paradata, which refers to meta-information, data, or, in a broader sense, things that are potentially informative of practices (Huvila, 2022, 2025). The intersections and differences between information literacy-as-practice and paradata provide a useful starting point to enquire more closely into the so far relatively little studied nexus of *information in social practices* (Cox, 2013), paradata (or, meta-information) *on* practices and the practices themselves. By doing this, it is possible to elicit new understanding of practices, literacies, information, paradata and their linkages.

The aim of this paper is to advance conceptual understanding of the nexus of literacies and paradata. Drawing on an interview study with professionals working with archaeological data, and focusing on their views on the premises of data literacy, i.e. competent sociocultural data practice in archaeology, we inquire into (RQ1) what it takes to be(come) literate, and (RQ2) how paradata functions in the context of literacy.

Literature review

Literacies

In contrast to our practice-based approach, much of the earlier literature on data-related literacies adopts a skills- and competencies-based approach. Fernández-Pascual and colleagues (2024) argue that even if discipline-specific data literacy concepts have emerged, data literacy is often conceptualised as instrumental – either data analysis or lifecycle oriented – competence, and contrasted to subject expertise. In addition, Bowler and Shaw's (2024) recent review suggests that while interest in data literacy is growing, the majority of studies are conducted in relation to formal learning, digital (big) data and everyday life. The notion of data literacy has also been introduced in the conceptual vocabulary of archaeology, with a focus on highlighting the need for

archaeologists to understand data both conceptually and in practice in the context of archaeological practice (Kansa and Kansa, 2021; also e.g., Garstki, 2022).

We argue that such one-size-fits all approaches to data literacy have led to a problematic narrowing of the definitions of both data and literacy. Our ongoing research into the information practices of archaeologists, including data creation and re-use, which includes both interviews and ethnographic fieldwork, has demonstrated that data in archaeological contexts can take on many forms spanning many specialties and subdisciplines, and can often draw its authority from the personal, embodied knowledge and expertise of members of the excavation team. Because of this, we argue that if we want to truly understand the complexity of data creation and data literacy within such fields, then a broader, practice-based approach is needed. We also argue that in the context of such a socially-contextualised understanding of data, the concept of paradata becomes an invaluable conceptual tool.

Paradata

While a breadth of paradata definitions and use cases exist across information science (see e.g., Huvila et al., 2024b; Sköld et al., 2022) and for example archaeology (Reilly et al., 2021) and survey research (Sakshaug and Struminskaya, 2023), paradata commonly refers to the diverse means of acquiring, developing and sharing understanding of past computational and human doings (activities, operations, decisions, deliberations) and their settings (social organisation, material conditions, intellectual horizons) of a data product. From a practice-led perspective, paradata is generally understood as (meta) information, or data, about practices, or more specifically things and occurrences that can be appropriated as informative of them (Sköld et al., 2022). Earlier studies also show that paradata during the course of research work is identified and enacted drawing on diverse sources often including narrative descriptions of practices, formal metadata, visualisations, photographs, audio and video recordings (Sköld and Andersson, 2025).

Although it is established in the comparably large literature on paradata needs that the principal use of paradata is as cues to gain insights regarding past machine or user data interactions in order to gauge data trustworthiness and usefulness (Davet et al., 2022; Huvila, 2020; Yakel et al., 2024), there is little research delving into how paradata functions in relation to how abilities are learned and competencies are gained, and what it means to be a competent practitioner. What has been found, however, is that paradata and paradata usefulness are malleable and emerge in different ways across disciplines and settings of work. Börjesson et al. (2022) show that paradata can exist in the minutiae of data organisation and description in datasets, requiring a significant amount of understanding of how the data was created and managed in its original context of creation to be discernible. Paradata can emerge from collectively enacted data practices for both junior and senior learners, requiring presence in a particular time and space to become possible to operationalise (Huvila and Sköld, 2021). The literature also underlines that there are technical (understanding hardware settings; Gant and Reilly, 2018), epistemic (knowing the research method employed; Huvila, 2020; see also Baker and Yarmey, 2009), knowledge-organisational (being able to see how a dataset has been structured; Börjesson, Sköld, et al., 2022) thresholds that have to be surpassed before paradata can be identified and used. Sufficient domain and contextual knowledge here emerges as a key component in being able to discern what paradata is available and to gauge its applicability in supporting the task at hand.

Methods and material

The findings of this paper are based on semi-structured interviews (N=33; referred to as A-AG in the analysis) with data professionals in the field of archaeology (Börjesson and Sköld, 2021). The interviews were carried out between 2020 and 2021 and focused on what capacities are required to be literate in information practices (RQ1) and better understanding the impacts of paradata in this setting (RQ2). There was also a data biographical element in the interview approach (cf.

Thomer et al., 2018) that sought to operationalise the intersections of paradata and data literacies by interrogating the data interaction histories of data and datasets the participants had created, curated or otherwise used in recent times. Interviews lasted 60-120 minutes and were conducted in Swedish and English depending on participant preferences. Translations of quotes into English were done by the authors. Covid-19 restrictions necessitated that all interviews were conducted and recorded using videoconferencing applications. A professional transcription service was employed to transcribe the interviews in clean verbatim. The interview transcripts were analysed using a method inspired by the constant comparative method and close reading to iteratively identify themes and generate analytical narratives to describe them.

The interview participants worked in a wide range of data-centred roles and capacities in the archaeological domain, principally as researchers and data managers. Participant recruitment followed a purposive sampling strategy (Robinson, 2014) with the aim to as broadly as possible represent different archaeological subdisciplines, career stages and archaeological data work expertises. A limitation of this sampling strategy is that the findings encompass only the subsets of archaeological data work encompassed in the study. An advantage of the heterogeneous sample is that it reflects the diverse epistemologies and data practices present in archaeological datawork, and offers favourable conditions for the study of grasping the tangled array of interactions between literacies-in-practice and paradata in the field.

Analysis

In the analysis, following the two research questions of the study, we identified facets of being and becoming literate (RQ1) in working with archaeological data, and investigated the role of paradata in the process (RQ2).

Facets of being literate (RQ1)

The analysis points to four overlapping senses of being literate with archaeological data (Figure 1). The interviewees described the necessity of 1) being literate in the archaeological practice that underpins the generation of archaeological data (*information making literacy*), to have at least rudimentary 2) literacy in the practices of data(base) design and management (*syntactic literacy*), and being 3) literate with the practices pertaining to planned data use (*information taking literacy*). In addition, the interviewees underlined the need to be 4) literate with the information (here, archaeological data) itself in an ontological sense (*interpretative literacy*). It seems that the interviewees with data management related work duties were somewhat more inclined to discuss information making and syntactic literacies in comparison to those primarily working as researchers.

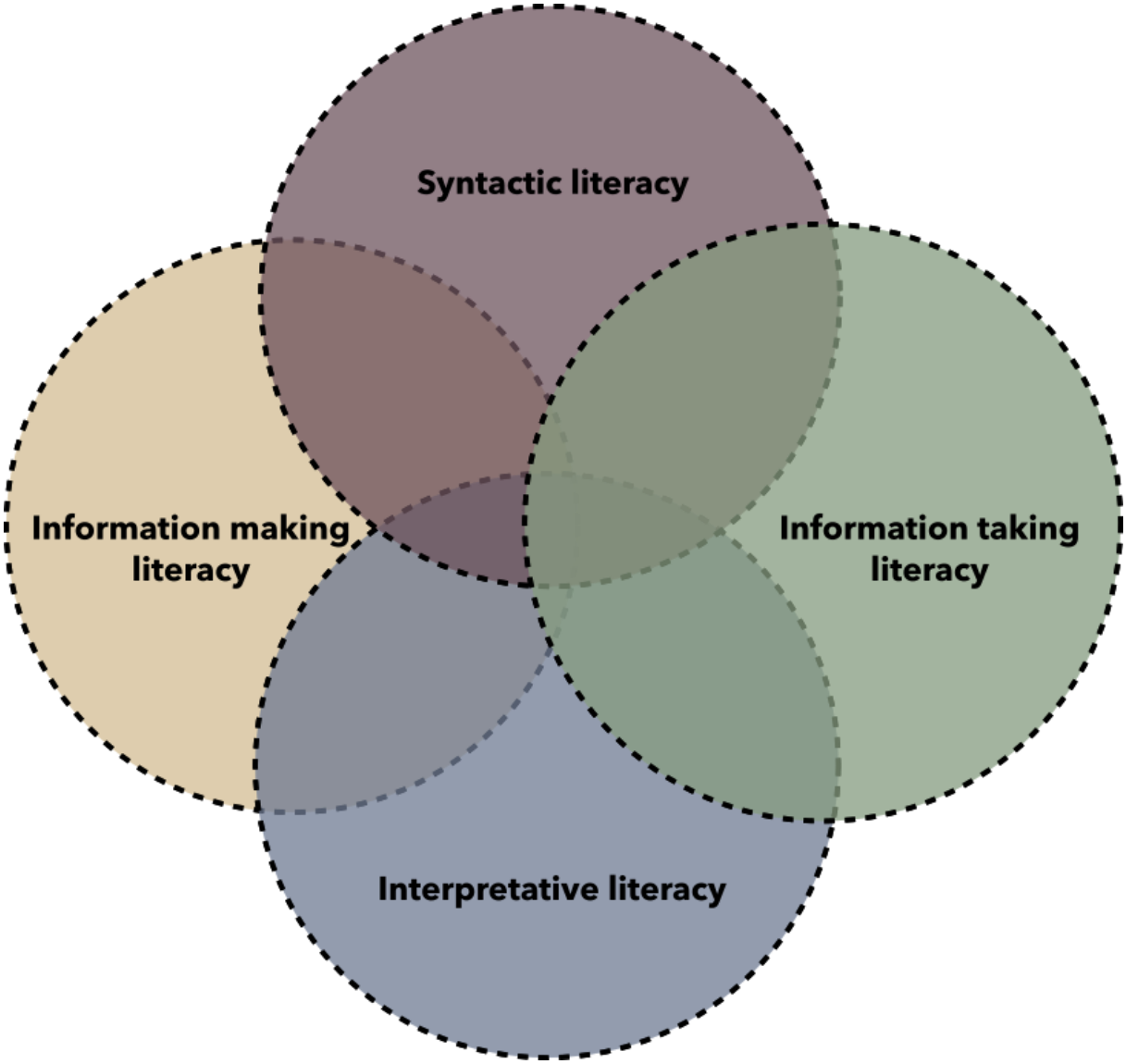

Figure 1. Four facets of being literate with archaeological data.

The analysis suggests further that the four senses of being literate form together a set of foundational facets of what constitutes *competent practice* with archaeological data. They all operate in parallel (e.g., A, C). For example, a field archaeologist who makes and documents observations during an excavation or survey needs to be proficient in both archaeological practices of information making but also at least to a certain extent in the syntactic practice of turning observations into a particular type of *data* (A, C, P). Similarly to being competent, it is equally crucial for literacy to understand what competences one is lacking (K) and how they differ from each other (G).

Information making literacy

A critical premise for working with archaeological data according to the interviewees is to understand archaeological practice and its relation to data being generated i.e. being literate with different means on archaeological practice of creating data in different countries, projects, contexts (C, I, AC). This broader competence, which we refer to as *information making literacy*, involves mastering the epistemics of how information making works within a particular field—

whether through observation, measurement or inference. It is about knowing '*the vocabulary, like the technical words of the discipline*' (A, also Z), and in archaeology, knowing research methodology (A), how archaeologists identify finds and features (R) and how they document (AC) and create plans and drawings (A). It encompasses an understanding of to what degree and in what terms a particular information generating practice is characterised by diversity and individualism or standardisation (Z).

Syntactic literacy

In parallel to understanding archaeology, the interviewees highlighted the need of at least rudimentary literacy in the technicalities of data management, database design and the data episteme i.e. understanding of how observations are transposed into *data* (G, I, K, M). In a broader sense, this competence we term *syntactic literacy* is about understanding the practices of how information is constructed and structured from its ingredients (e.g., observations, measurements, words, experiences, feelings and their combinations through *representational exchange*, see Huvila, 2024) to become information in a specific episteme.

Within the currently dominant episteme of digital data practice in archaeology, syntactic literacy encompasses a command of how to define and use formal vocabularies (Z) and computational ontologies (Z), create datasets (I) and understanding how databases and database systems work (G, R) and differ from spreadsheets (R). Archaeologists working with digital spatial data are also expected to be syntactically literate with geographical information systems and map-making (A, P).

Despite the dominance of digital data practices, the interviewees made comments on non-digital syntactic literacies of, for example, keeping a diary, filling out a form and drawing a map. Independent of the specific technologies used, an information maker needs to know how to '*systematise*' (I) and manage (O) it according to a particular episteme. While the use of spreadsheets was deemed substandard from the perspective of state-of-the-art digital syntactic literacy (R), it was also considered to be a baseline for what can be expected of an average archaeologist (AC). Interviewee (R) described the use of spreadsheets not as data-oriented but rather as a form of visual practice.

Information taking literacy

The interviewees also suggested a need to be literate in the practices of taking archaeological data into use in particular ways. A key aspect of what we call *information taking literacy* is an insight that particular types of information (e.g., research data, drawings, photographs or handwritten notes) are possible and relevant to take into use in specific practices in different ways. '*There's also kind of a skill to it*' as interviewee C noted that differs from the literacy of making information and data, for example, creating and using databases. Being a literate user and developer of archaeological databases is not the same (G). Interviewee I underlined further how the necessary type of mastery of archaeological practice depends on the type of data practice pursued. Part of it is also to understand the complexities of archaeological data and practices and their implications to using data (K).

Interpretative literacy

Finally, the interviewees underlined the importance of understanding and knowing archaeological data. *Interpretative literacy* is what we term as being literate in the content of information. It is the starting point of source criticism from the perspective of knowing the subject matter (I). With archaeological data, in contrast to understanding how a plan is drawn (cf. information making literacy), it is about understanding what the '*contents of the plan represent*' (A) and more generally, on what (cf. how in information making literacy) archaeologists focus when they are documenting and making observations (AC). Interviewee I described it in terms of a '*cultural historical competence*', referring to archaeology as a cultural historical discipline.

Becoming literate (RQ1)

Besides facets of what it takes to *be* literate, we also identified factors that underpin *becoming* literate with archaeological data (summarised in Figure 2).

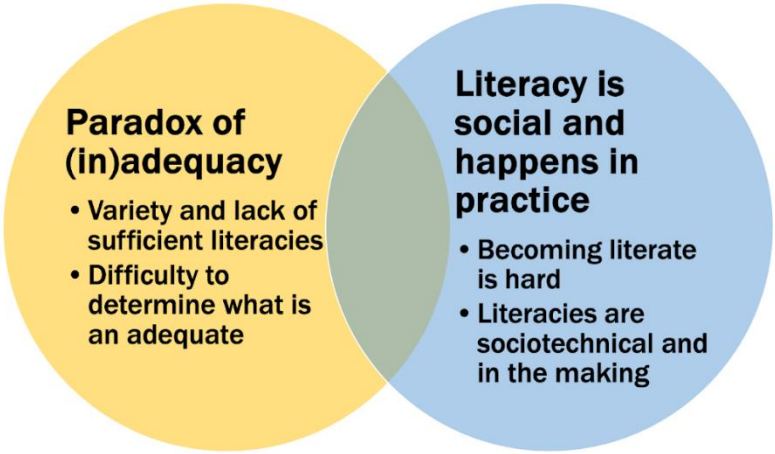

Figure 2. Factors that underpin becoming literate with archaeological data.

Paradox of (in)adequacy

The interplay and concurrency of the literacy facets means that a *competent enough* practice with archaeological data is a moving target. The adequacy of particular literacy depends on the available competences in others together with the aims of working with data.

First, several interviewees commented on the *variety and lack of sufficient literacies among archaeologists*, mostly pertaining to lack of syntactic competence of using digital tools and understanding how data and databases work as means for representing archaeological observations (A, C, G, I, K, R), but also in archaeological information making literacy of, for example, identifying specific types of artefacts (R). Both syntactic literacy (A, AC) and archaeological practice-related information making literacy (K) was also criticised for being unevenly distributed among archaeologists, sometimes leading to variance and inconsistencies in archaeological data (O). It is problematic '*[i]f [archaeologists] lack some of the skills*' (I) and if there is a gap, for example, '*between the technical expertise and the practical skills of the archaeologist and also the knowledge of the data itself, that gap is quite substantial*' (K).

Everyone was considered to need to learn basic data management skills (I, K, P). At the same time, it was considered that sophisticated digital skills were beyond what could be expected or needed (P). Working properly with data was considered to require a specialist (K, O) or a special organisational unit or similar that works on '*another level*' with the data (I) than an average archaeologist.

Second, a related problem to perceived lack and imbalance of competences was the *difficulty to determine what is an adequate level of specific literacies*. Partly, the degree of literacy is impossible to measure and communicate in detail (R). As interviewee P suggested, it may be possible to say that some are skilled and some are not, but who is competent and how differs considerably (P). In archaeology, sometimes students with a workable level of interpretative literacy and theoretical

competence in archaeological information making act as supervisors to local workers with decades of practical experience of archaeological information making (P). The type and time of education also affect competences and ways of seeing things (Y). People also have different subjective understandings of their level of competence, and it is difficult to say how sure they are about their verdicts (Z).

Literacy is social and happens in practice

The paradox of (in)adequacy in literacies is both a result of and mitigated by the fact that they are interlinked, depending on(O) and compensating for each other, are social and they are learnt and become (or happen) in practice.

First, *becoming literate is hard.* Developing literacies requires time and effort (V) some have and others lack (K, O, R). Training can be helpful and facilitate both information making and taking (R). The lack of formal training especially in data management can be problematic in how resulting insufficient competence makes working time-consuming and difficult, and leads to *data* that is not adequate in the dominant episteme of digital data or any other syntactic regime beyond the data-maker's own (C, V, AB, AC). This has consequences on the degree of information taking literacy required for '*tinkering*' and making sense of the data (C, AB). '*It's important to understand ... that [data is] not ... necessarily derived in quite the same way so you have to ... know, you have to understand that*' (C). It turns into a serious problem if data is formally correct but does not correspond with the observation (C, AC).

Variety and learning by doing are, however, also a hallmark of archaeological practice that needs to cope with an extreme variety of information sources and potentially relevant perspectives (C, I, V, AB, AD). Much of the variety can be explained by the fact that, while '*everyone's trying to do their best ... people learn from their lab or from a friend or from an advisor or whatever, and everyone does things in a different way*' (R), but also in that it is extremely difficult, if possible at all, to converge information making literacy and syntactic literacy in terms of common terminology and classifications (AD). From the perspective of a particular practice on archaeological information making, it can be useful to refuse to subscribe to syntactic literacy based on the digital episteme, and, for example, use a spreadsheet in a manner that is epistemically useful in that practice even if would make it impossible to directly extract the data to a database (R).

Second, *literacies are sociotechnical and in the making*. They are sourced from the sociotechnical sphere and intertwined in the social fabric of archaeological data practice. A common strategy to compensate for a lack of competence is to rely on social contact (C, O, AG). Sometimes a student in computer science can provide a necessary complement to the required syntactic literacy to solve a problem to a point that others can do it without exact knowledge of all technical details (K). Besides people, literacies can also be outsourced to technologies, for example, computer programs and structured data input forms (O). However, converging across the sociotechnical sphere requires trust (I, K) and shared understanding of each other's perspective: '*it is better to work with people that you know ... than having the perfect team with the perfect skills*' (AE).

Paradata and literacies (RQ2)

In the interview record, multiple points of convergence between literacies and paradata can be identified. As a whole, paradata emerges as the cues that help people to understand the practice related to archaeological information making and taking, tools and epistemes.

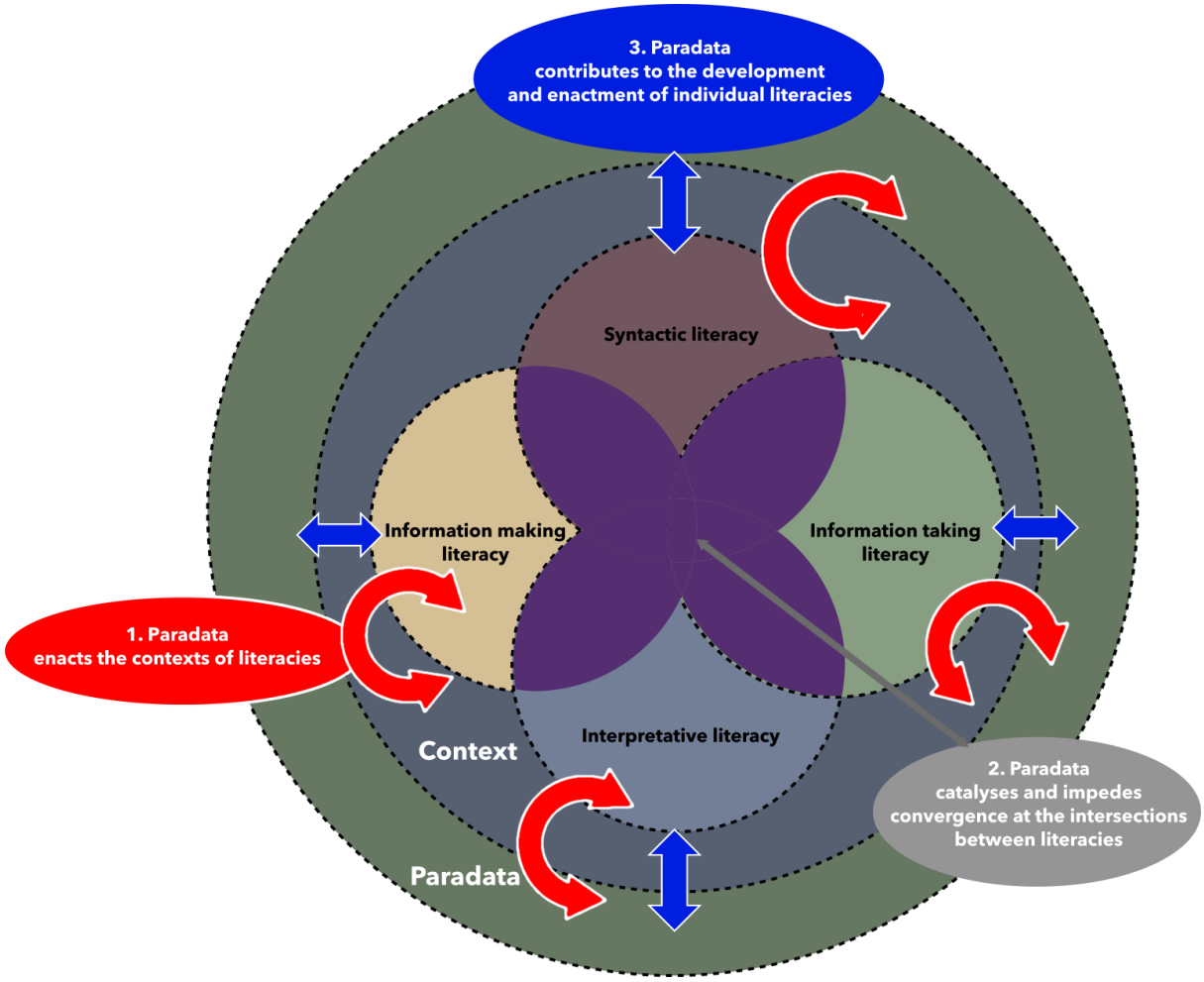

Figure 3. How paradata functions in the context of literacy.

First, paradata can *contribute to development and enactment of individual literacies* (red in Figure 3). It can contribute to knowing data (interpretative literacy) by helping to understand how it came into being (A, K), to information taking literacy through explaining the complexity (K), practices and criteria of its making (I, K) and validation (O), and pointing to how it might be used (M). Paradata could also contribute to syntactic literacy; for instance, by conveying how a database works (cf. A, G) and to information making literacy by shedding light on the evolution of the archaeological practice (cf. A, I), background, practices and language of individuals engaging in the making of archaeological data (A, C, I, R, Y, Z).

The observation above that individual literacies are simultaneously inadequate and compensate for each other, which means that a precise mapping of paradata types to literacies is not possible. There are, however, some observable patterns, summarised in Table 1.

Literacy	Paradata type	Examples in the interview record	Example formats
Information making literacy	Knowledge organisation and representation paradata (Börjesson, Sköld, et al., 2022); Knowledge making paradata (Börjesson, Huvila, et al., 2022).	Currently applicable methods (C, D, F, H, I); Information on standards (C, E, H, I) and formats (A, F) used.	Formal standards.
Information taking literacy	Knowledge making paradata (Börjesson, Huvila, et al., 2022); Methods paradata; Provenance paradata (Börjesson, Sköld, et al., 2022).	Nature and context of, and rationale underpinning data (C, G, H, I); epistemological culture (C).	Narratives of data creation practices.
Syntactic literacy	Knowledge organisation paradata (Börjesson, Huvila, et al., 2022); Knowledge organisation and representation paradata; Methods paradata (Börjesson, Sköld, et al., 2022).	Data codes and formats (A, D, H); Data structures and classifications (C, D, F); Instruments (F).	Templates; Training and apprenticeship; Workflows, Procedures; Instruction manuals.
Interpretative literacy	Knowledge making paradata (Börjesson, Huvila, et al., 2022); Scope paradata; Methods paradata; Provenance paradata (Börjesson, Sköld, et al., 2022).	Data coverage: spatial, temporal (B, C, E, G, I); Detail (G): Nature and context of, and rationale underpinning data (C, G, H, I).	Explanations of the rationales and purposes of data creation; Examples of data.

Table 1. Literacy facets, their associated paradata types and formats.

Second, *paradata can catalyse and impede convergence at the intersections between literacies* (grey in Figure 3). Paradata can convey crucial aspects of, for example, syntactic literacy to those creating archaeological data (K), and mitigate the difficulty of working with others lacking a particular literacy (I). There are, however, also situations when the work with paradata and the clarity of documentation are not prioritised or it is kept difficult to use on purpose:

> *I think a lot of the reluctance is tied to feelings about prestige and the need to secure a career for yourself by maintaining a dataset for your exclusive use. And then, if that becomes available to other people, you may lose some of the priority and the opportunities that you believe you're entitled to.* (AD)

Third, the *paradata enacts the contexts of literacies* (blue in Figure 3). Partly, while the interview record suggests that the same information can contribute to the enactment of multiple literacies, it also points to the fact that not all forms of paradata are compatible with all varieties of literacies. For example, interviewee AC commended how papers accompanied by Jupyter notebooks are good examples of how data and workflows can be brought and kept together but doubted at the same time whether they are technically too complicated for the average archaeologist at the moment, and pondered about ways how to achieve something similar that would be compatible with the practices of making and using archaeological data. Through being inscribed and embodied in diverse physical and conceptual forms, paradata enacts the contexts and contextualities of literacies, making them visible and practicable in context.

Discussion

What it takes to be literate

The findings underline the complexity and contextuality-in-practice of how becoming and being literate with archaeological data is enacted. Rather than comprising a single practice of becoming and being data literate, the findings suggest the presence of parallel literacies of 1) making information according to the conventions of archaeological practice (information making literacy), 2) taking such information into use (information taking literacy), 3) understanding the technical practices of how information is constructed and structured from its ingredients using a particular set of techniques and tools (syntactic literacy), and 4) interpreting and understanding the information and its meaning (interpretative literacy). They align with sociomaterial conceptualisations of engaging with literacies (e.g. Lloyd, 2010; Pourbaix, 2000) and classic descriptions of what constitutes a competent practice (e.g., Law and Lynch, 1988), however, providing a systematic enumeration of its key facets. The facets differentiate technical and subject-oriented trajectories of expertise typically separated in the literature (Fernández-Pascual et al., 2024), however nuancing the picture by distinguishing information making, interpretation and taking as distinct dimensions of the latter. Syntactic literacy comes closest to being a technical form of infrastructural literacy (not to mix it up with literacies of infrastructure, e.g. Gray et al., 2018) as the practice of mastering the underpinning technical and epistemic means of dealing with data in particular ways. However, similarly to other identified facets, it is not operating in isolation, being linked both to the epistemic underpinnings of specific data-technologies and how they are applied in context.

This means also that the different literacies as facets of being literate complement and compensate for each other but also contradict in their expectations and emphases on what kinds of data are central to upholding archaeological practice. This makes literacies a moving target, simultaneously adequate and inadequate. We suggest that previously observed diverging preferences of paradata and other forms of data documentation between data creators, managers and users can be explained in terms of their addressing of the different literacy facets. It appears that problems and solutions in data work are contextual to specific literacies, being based on assumptions and modes of operation that sometimes fail to generate or leverage data that is fully compatible with those of other literacies (e.g., Börjesson, Huvila, et al., 2022; Huvila et al., 2024a; Olsson et al., 2024). This was particularly obvious between the information making and syntactic literacies versus information taking and interpretative literacies. The first-mentioned emphasise formal and technical knowledge representation (and organisation) paradata, whereas the latter foreground descriptive and contextual types of scope and provenance paradata (Börjesson, Huvila, et al., 2022; Börjesson, Sköld, et al., 2022). Data makers' engagement with information making and syntactic literacies and data users on information taking and interpretative literacies further explain the differences in their respective wants and needs regarding structured versus narrative data documentation (Huvila et al., 2024a).

The findings on the complexity of literacies point further to the occasional productivity of *unliteracy* (Huvila, 2018; cf. Plaut, 2023) – of refusing to subscribe to a particular literacy practice but instead of relying on others, not knowing or documenting everything. From the perspective of a particular practice on archaeological information making, it can be useful to refuse to subscribe to syntactic literacy based on the digital episteme and, for example, use a spreadsheet in a manner which makes it impossible to directly extract the data to a database using it according to the premises of a visual rather than datafied syntactic literacy, simultaneously emphasising a different information taking and interpretative practice to structured information making (as, for example, for interviewee R).

How paradata functions in the context of literacy

In the interview record, getting and staying informed about a competent practice through engaging with paradata serves three major purposes. By informing about specific aspects of practices, it can help to develop and enact individual data-related literacies. Paradata is clearly not embodied competence made explicit but rather an ingredient of learning something akin to craftsman's abstractions (Høgseth, 2012) in how they encapsulate conceptualisations of the practice. Rather than functioning as a substitute, a broader understanding of a practice as a whole facilitated by paradata can compensate for a lack of mastery of some literacies and contribute to the convergence in the broader *landscape* (cf. Lloyd, 2010) of literacies. Further, through being informative about practices, paradata enacts contexts of – in our study archaeological – data practice resulting in better understandings of data and how and where it can work, i.e., aspects that have repeatedly been reported to be poorly documented and difficult to grasp in shared datasets (e.g., Faniel et al., 2019; Kim and Yoon, 2017).

However, rather than merely informing about processes or doings in a narrow technical sense, the findings demonstrate how paradata through informing on '*how things are done*' conveys understanding of not only practices but also their associated practice architectures (Kemmis, 2019; Olsson et al., 2024), and broader habitat that Stengers (2005) terms their ecology. From this perspective, we posit that many of the aspirations of what '*data literacy*' (cf. Kansa and Kansa, 2021) should achieve are less concerned about mastering data itself than paradata (see also Huvila et al., 2024). When data is recognised as a process rather than an immutable thing, the nucleus of data literacy is understanding historical, contemporary and future practices that influence(d) data, in effect mastering what we term paradata. The technical exercise of grasping data and its technical lifecycle is only part of being literate. Similarly to how understanding written work – a literary piece or an everyday text – requires competence in understanding it in its context situated in the process of its becoming beyond deciphering it as a sequence of words, being literate with data requires a similar grasp of it in its full, constantly evolving complexity.

A theoretical benefit of discussing paradata and literacies together is in how they both underline the importance of the contextuality of practices. While context has been debated in information studies for decades, resulting in a plethora of models, it still to a large extent lacks functional operationalisations (cf. Agarwal, 2017; Faniel et al., 2019). In contrast to vaguely suggesting that there is a need for contextual information, paradata offers a way of approaching context in a systematic manner through explicating mechanisms of how specific types of information-on-practice link to competent practice-in-context.

Interestingly enough, the importance of such competent practising of the contextuality of work through developing literacy through generation of what essentially is paradata has been discussed in the empirical context of this study – archaeology – from the 1990s onwards in the context of what is known as *reflexive archaeology* (Hodder, 1997; also, Berggren et al., 2015). The aim of the approach is to encourage archaeologists to reflect on their field practice in real time and write down reflections and interpretations of the practice – in practice, paradata – to avoid unreflective collecting of contextually detached *data*. Studies enquiring into reflexive archaeology suggest that producing reflexive paradata requires methods that are contextually relevant to a specific archaeological investigation (Sandoval, 2021), and good understanding of its context together with '*experience and skill*' (Sandoval, 2020). Hardly surprisingly, these measures map onto the identified literacy facets relevant to making and taking archaeological data and knowledge.

Conclusions

A key insight from this study is how paradata, as a concept referring to practice-related meta-information, provides means to operationalise what can be known and communicated about practices as a central contextual constituent of being and becoming literate. The findings nuance

earlier propositions of the potential importance of paradata in the context of data literacies. The present enquiry into archaeological data practice shows further how being literate is composed of multiple facets (literacies) that are in flux, distinct but overlapping, aligned and conflicting. A particularly blatant breach was identified between the literacies of making and understanding syntactic means of organising information and taking and understanding it in context. Becoming and being literate as a whole entail navigating them, not necessarily mastering, but learning enough of them all for staying afloat. Different types of information – or paradata – are helpful in this process, with specific types of information having affinities to enacting particular types of literacies, contexts and convergences in between. Finally, the exploration in archaeological data practice showed that being literate was more of a question of being literate in working with paradata about archaeological data practice than mastering the data itself, essentially pointing to the fact that data literacy, and literacies in general, could better be framed as literacies of paradata.

Acknowledgements

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement No 818210 as a part of the project CApturing Paradata for documenTing data creation and Use for the REsearch of the future (CAPTURE).

About the author(s)

Dr Isto Huvila is Professor of Information Studies at the Department of ALM at Uppsala University in Sweden. His primary areas of research include information and knowledge management, information work, knowledge organisation, documentation and social and participatory information practices. He can be contacted at isto.huvila@abm.uu.se.

Dr Michael Olsson is a researcher at the Department of ALM at Uppsala University attached to the CAPTURE project. He is an active researcher in the field of information practices research and his work is essentially interdisciplinary in nature. He can be contacted at michael.olsson@abm.uu.se.

Dr Jessica Kaiser is a Visiting Lecturer in Egyptian Archaeology at the Department of Archaeology and Ancient History and Conservation at Uppsala University. She was previously a post-doctoral researcher with the CAPTURE project. Her current research focuses on research data management and the history of cultural heritage management in archaeology. She can be contacted at jessica.kaiser@arkeologi.uu.se

Dr Olle Sköld is a Senior Lecturer in the Department of ALM and the director of Uppsala University's Master's Programme in Digital Humanities. His research is characterised by a broad interest in the ALM field, research data creation and use and digital humanities. He can be contacted at olle.skold@abm.uu.se.

Dr Lisa Andersson has worked as a researcher in the Department of ALM at Uppsala University in Sweden. Her research focuses on research information, including research information management systems, data descriptions, data publishing and use. She can be contacted at lisa.andersson@ki.se

References

Agarwal, N. K. (2017). *Exploring context in information behavior: seeker, situation, surroundings, and shared identities*. Morgan & Claypool. https://doi.org/10.2200/s00807ed1v01y201710icr061

Baker, K. S., & Yarmey, L. (2009). Data stewardship: environmental data curation and a web-of-repositories. *The International Journal of Digital Curation*, 4(2), 12–27.

Berggren, Å., Dell'Unto, N., Forte, M., Haddow, S., Hodder, I., Issavi, J., Lercari, N., Mazzuccato, C., Mickel, A., & Taylor, J. (2015). Revisiting reflexive archaeology at Çatalhöyük: integrating digital and 3D technologies at the trowel's edge. *Antiquity*, 89(344), 433–448.

Börjesson, L., Huvila, I., & Sköld, O. (2022). Information needs on research data creation. *Information Research*, 27(Special Issue), isic2208. https://doi.org/10.47989/irisic2208

Börjesson, L., & Sköld, O. (2021). *The making and use of paradata: an interview study* [Dataset]. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-455730

Börjesson, L., Sköld, O., Friberg, Z., Löwenborg, D., Pálsson, G., & Huvila, I. (2022). Re-purposing excavation database content as paradata: an explorative analysis of paradata identification challenges and opportunities. *KULA: Knowledge Creation, Dissemination, and Preservation Studies*, 6(3), Article 3. https://doi.org/10.18357/kula.221

Bowler, L., & Shaw, C. (2024). Trends in data literacy, 2018-2023: a review of the literature. *Information Research an International Electronic Journal*, 29(2), Article 2. https://doi.org/10.47989/ir292822

Cox, A. M. (2013). Information in social practice: a practice approach to understanding information activities in personal photography. *Journal of Information Science*, 39(1), 61–72. https://doi.org/10.1177/0165551512469767

Davet, J. E., Hamidzadeh, B., Franks, P. C., Bunn, J., Hamidzadeh, B., Franks, P. C., & Bunn, J. (2022). Tracking the functions of AI as paradata & pursuing archival accountability. *Archiving Conference*, 19, 83–88. https://doi.org/10.2352/issn.2168-3204.2022.19.1.17

Faniel, I. M., Frank, R. D., & Yakel, E. (2019). Context from the data reuser's point of view. *Journal of Documentation*, 75(6), 1274–1297. https://doi.org/10.1108/JD-08-2018-0133

Fernández-Pascual, R., Pinto, M., & García Marco, F. J. (2024). Emergence and evolution of data literacy: Insights from a bibliometric study. *Journal of Librarianship and Information Science*, 09610006241265102. https://doi.org/10.1177/09610006241265102

Gant, S., & Reilly, P. (2018). Different expressions of the same mode: a recent dialogue between archaeological and contemporary drawing practices. *Journal of Visual Art Practice*, 17(1), 100–120. https://doi.org/10.1080/14702029.2017.1384974

Garstki, K. (2022). Teaching for data reuse and working toward digital literacy in archaeology. *Advances in Archaeological Practice*, 10(2), 177–186. https://doi.org/10.1017/aap.2022.3

Gee, J. P. (2000). The new literacy studies: from 'socially situated' to the work of the social. In D. Barton, M. Hamilton, R. Ivanič, J. P. Gee, K. Jones, J. Maybin, F. Ormerod, S. Pardoe, K. Pitt, R. de Pourbaix, K. Tusting, A. Wilson, & M. Hamilton (Eds.), *Situated literacies: reading and writing in context* (pp. 177–193). Routledge.

Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society*, 5(2), 2053951718786316. https://doi.org/10.1177/2053951718786316

Hodder, I. (1997). 'Always momentary, fluid and flexible': towards a reflexive excavation methodology. *Antiquity*, 71(273), 691–700. https://doi.org/10.1017/s0003598x00085410

Høgseth, H. B. (2012). Knowledge transfer: the craftsmen's abstraction. In W. Wendrich (Ed.), *Archaeology and apprenticeship: body knowledge, identity, and communities of practice* (pp. 61–78). University of Arizona Press.

Huvila, I. (2018). Alternatives to being information literate. In S. Kurbanoğlu, J. Boustany, S. Špiranec, E. Grassian, D. Mizrachi, & L. Roy (Eds.), *Information literacy in the workplace* (pp. 813–821). Springer. https://doi.org/10.1007/978-3-319-74334-9_82

Huvila, I. (2020). Information-making-related information needs and the credibility of information. *Information Research*, 25(4), paper isic2002. https://doi.org/10.47989/irisic2002

Huvila, I. (2022). Improving the usefulness of research data with better paradata. *Open Information Science*, 6(1), 28–48. https://doi.org/10.1515/opis-2022-0129

Huvila, I. (Ed.). (2025). *Paradata: documenting data creation, curation and use*. Cambridge University Press.

Huvila, I., Andersson, L., & Sköld, O. (2024a). Patterns in paradata preferences among the makers and reusers of archaeological data. *Data and Information Management*. https://doi.org/10.1016/j.dim.2024.100077

Huvila, I., Andersson, L., & Sköld, O. (Eds.). (2024b). *Perspectives on paradata: research and practice of documenting data processes*. Springer. https://doi.org/10.1007/978-3-031-53946-6

Huvila, I., & Sköld, O. (2021). Choreographies of making archaeological data. *Open Archaeology*, 7(1), 1602–1617. https://doi.org/10.1515/opar-2020-0212

Huvila, I., Kaiser, J., Sköld, O., & Andersson, L. (2024). Paradata literacy and the challenges of research data management. *Informaatiotutkimus*, 43(3–4), 109–113. https://doi.org/10.23978/inf.148594

Kansa, E., & Kansa, S. W. (2021). Digital data and data literacy in archaeology now and in the new decade. *Advances in Archaeological Practice*, 9(1), 81–85. https://doi.org/10.1017/aap.2020.55

Kemmis, S. (2019). *A practice sensibility*. Springer. https://doi.org/10.1007/978-981-32-9539-1

Kim, Y., & Yoon, A. (2017). Scientists' data reuse behaviors: a multilevel analysis. *JASIST*, 68(12), 2709–2719. https://doi.org/10.1002/asi.23892

Law, J., & Lynch, M. (1988). Lists, field guides, and the descriptive organization of seeing: birdwatching as an exemplary observational activity. *Human Studies*, 11(2/3), 271–303.

Lloyd, A. (2010). *Information literacy landscapes: information literacy in education, workplace and everyday contexts*. Chandos.

Olsson, M., & Lloyd, A. (2017). Being in place: embodied information practices. *Information Research*, 22(1), paper 1601.

Olsson, M., Sköld, O., & Andersson, L. (2024). Layers upon layers: data sharing & reuse challenges in archaeological contexts. *Information Research an International Electronic Journal*, 29(2), Article 2. https://doi.org/10.47989/ir292840

Plaut, E. (2023). Strategic illiteracies: the long game of technology refusal and disconnection. *Communication Theory*, 33(1), 21–31. https://doi.org/10.1093/ct/qtac014

Pourbaix, R. de. (2000). Emergent literacy practices in an electronic community. In D. Barton, M. Hamilton, R. Ivanič, J. P. Gee, K. Jones, J. Maybin, F. Ormerod, S. Pardoe, K. Pitt, R. de

Pourbaix, K. Tusting, A. Wilson, & M. Hamilton (Eds.), *Situated literacies: reading and writing in context* (pp. 123–145). Routledge.

Reilly, P., Callery, S., Dawson, I., & Gant, S. (2021). Provenance illusions and elusive paradata: when archaeology and art/archaeological practice meets the phygital. *Open Archaeology*, 7(1), 454–481. https://doi.org/10.1515/opar-2020-0143

Robinson, O. C. (2014). Sampling in interview-based qualitative research: a theoretical and practical guide. *Qualitative Research in Psychology*, 11(1), 25–41. https://doi.org/10.1080/14780887.2013.801543

Sakshaug, J. W., & Struminskaya, B. (2023). Augmenting surveys with paradata, administrative data, and contextual data. *Public Opinion Quarterly*, 87(S1), 475–479. https://doi.org/10.1093/poq/nfad026

Sandoval, G. (2020). The relevance of Çatalhöyük reflexive diaries. In K. Kaercher, M. Arntz, N. Bomentre, X. L. Hermoso-Buxán, K. Kay, S. Ki, R. Macleod, H. Muñoz-Mojado, L. Timbrell, & I. Wisher (Eds.), *New frontiers in archaeology: proceedings of the Cambridge Annual Student Archaeology Conference 2019* (pp. 25–41). Archaeopress. https://doi.org/10.2307/j.ctv1dp0v4j

Sandoval, G. (2021). In pursuit of a reflexive recording. An epistemic analysis of excavation diaries from the Çatalhöyük research project. *Norwegian Archaeological Review*, 53(2), 135–153. https://doi.org/10.1080/00293652.2020.1854338

Savolainen, R. (2008). *Everyday information practices: a social phenomenological perspective*. Scarecrow Press.

Sköld, O. & Andersson, L. (2025). Paradata 'in the wild': how and where paradata emerges in research documentation. In *Paradata: documenting data creation, curation and use*. Cambridge University Press.

Sköld, O., Börjesson, L., & Huvila, I. (2022). Interrogating paradata. *Information Research. Proceedings of the 11th International Conference on Conceptions of Library and Information Science, Oslo Metropolitan University, 29 May – 1 June 2022*, 27(Special Issue), paper colis2206. https://doi.org/10.47989/colis2206

Stengers, I. (2005). Introductory notes on an ecology of practices. *Cultural Studies Review*, 11(1), 183–196.

Thomer, A. K., Wickett, K. M., Baker, K. S., Fouke, B. W., & Palmer, C. L. (2018). Documenting provenance in noncomputational workflows: research process models based on geobiology fieldwork in Yellowstone National Park. *JASIST*, 69(10), 1234–1245. https://doi.org/10.1002/asi.24039

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

Yakel, E., Faniel, I. M., & Robert Jr, L. P. (2024). An empirical examination of data reuser trust in a digital repository. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.24933