7

A Paradata Reference Model

Isto Huvila

7.1 Introduction

In this volume we have traversed from exploring how paradata has been discussed in the literature to exploring where it can be found and how it can be practised by data creators, users and curators. What has not been touched upon so far is how these different strings can be tied together to advance a conceptual understanding of paradata: what paradata is and perhaps even more importantly, what it does and in what kind of a space it operates. The purpose of this chapter is to delve into these questions and, through assembling and disassembling a reference model, to work towards a proposed theoretical synthesis of paradata and its workings. Before proceeding any further, it is fair to note as a word of warning to any casual readers that the aim of this chapter is to question, problematise and theorise rather than to provide practical advice.

Our synthesis needs, however, to start with its opposite, an ἀνάλυσις (analysis) of what paradata allows us to do. Building on how paradata was characterised in the introduction of this volume as a form of information relating to informational – and other varieties of – doings, a fruitful starting point is to think more closely about the implications of thinking in terms of doings. To this end, Ingold (2022) makes a useful distinction between making and doing. If we imagine we approach a researcher, or a construction worker as Ingold did, and ask what you do, the answer is likely to be rather different from that to the question of what you make. Answers about doing are likely to refer to tasks at hand (what is happening right now) whereas answers about making are more likely to relate to the final output of the process. For a researcher it could be something tangible, like a research article or the planned outcomes of

a research project, while for a construction worker it would relate to a building or a bridge depending on what they are building at the time.

Paradata strikes into this very distinction. In research work we are also often very much focused on what we are *making* and how to articulate and document it rather than what we are *doing* to achieve it. Rather than focusing on the outcome, we propose that paradata can help us to focus on doing. As '[d]oing ... is hidden inside the box of making' (Ingold, 2022, p. 224). While paradata may not be able to take 'doing' out of the box, it does afford the opportunity to at least open the lid and provide a glimpse inside.

The difficulty with paradata and what we, in this chapter, term working knowledge is that there is no one obvious perfect approach to opening the box. Articulating something that is already outside of it is much easier. This is demonstrated vividly in studies of such colloquial impromptu forms of documentation as collaborative tagging systems (Golder and Huberman, 2006) where specific systematic modus operandi are not enforced.

The word paradata itself alludes to the possibility to understand practices and processes in (new) ways that make them more manageable and countable, turning them into what Latour (1987) would call immutable mobiles. The ideal of making processes comprehensible through documentation has multiple parallels in contemporary information cultures. It ties into what has been described as the explainability turn (Berry, 2023), an aspiration to mitigate the apparent opaqueness of algorithmic processes and increase their transparency through providing explanations.

More than a desire to satisfy curiosity, having an explanation is increasingly regarded as a legal right in modern legislation. An explanation becomes a product, a category of things, in the same way that the documentation and comprehensibility of data-related practices and processes are increasingly treated in the literature (cf. Chapter 2) as a manufactured product that reifies an obligation. However, as the previous chapters in this volume have demonstrated, paradata is also distinctly doing the opposite. There is no one way of capturing or generating paradata, or a single explanation it is capable of enacting. Attempting to use paradata to make practices and processes countable, makes visible their uncountability, precariousness and instability.

7.2 The Model

Instead of starting with an explanation of a set of premises, here the story begins with the model itself before delving into discussing its details. The paradata reference model (Figure 7.1) frames paradata as related to practices

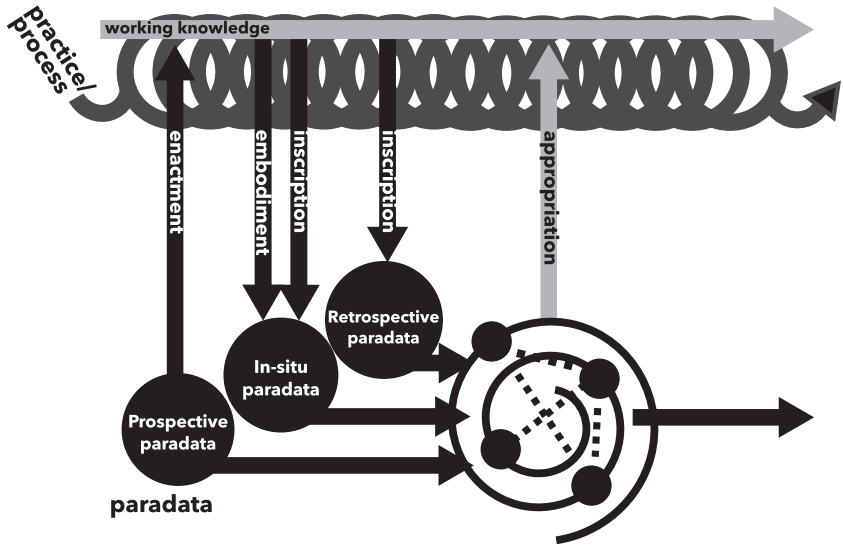

Figure 7.1 The paradata reference model.

and processes. *Working knowledge* on *practices and processes* (i.e. how practices and processes are known) and the practices and processes themselves engender paradata through *embodiment* and acts of *inscription*. Paradata turns back to working knowledge through *appropriation*. *Enactment* turns paradata to practices and processes. For *prospective paradata* engendered before its referent practice or process its enactment is also the very moment when it turns from a potential to actual paradata. Paradata is processual in the sense that it is perpetually *in the making* on a continuum and actualises as instances only through *temporal intersections* with the practices and processes it converges.

In parallel to being processual in time, consisting of interconnected types of information linked to each other, paradata forms a network. As the previous chapters have shown, while sometimes a single item might be enough to convey the necessary understanding of pertinent facets of a practice or process, the richness of paradata grows through assemblages where the different constituent parts come together to provide a richer understanding which goes beyond being simply the sum of its component pieces.

At this point, we turn our attention to the constituent parts of the model and return to discussing paradata as embodied or inscribed forms of working knowledge. We continue to explore what it means that paradata is created on a continuum where it is actualised through temporal intersections between practices and processes. This is followed by a closer look at the enactment of

paradata in practices and processes; projected and unintended outcomes of working with paradata; and finally inquiring into how paradata resides in the borderlands between practices and processes; and embedded and explicable through ways of knowing.

7.3 Working Knowledge and Paradata

We start by zooming into the centre of the model to examine a section (Figure 7.2) where *working knowledge* intersects with paradata in the nexus of what belongs to the realm of practices and processes (upper part of Figure 7.1, marked in light grey) and the realm of paradata and documentation (lower part, in black).

The reason for starting with working knowledge is that much of the knowledge people act upon on a daily basis is incorporated into the processes and practices themselves. It is infrastructural: taken for granted, inherited background knowledge (Tsoukas, 2013; Wittgenstein, 1969). It is not knowledge

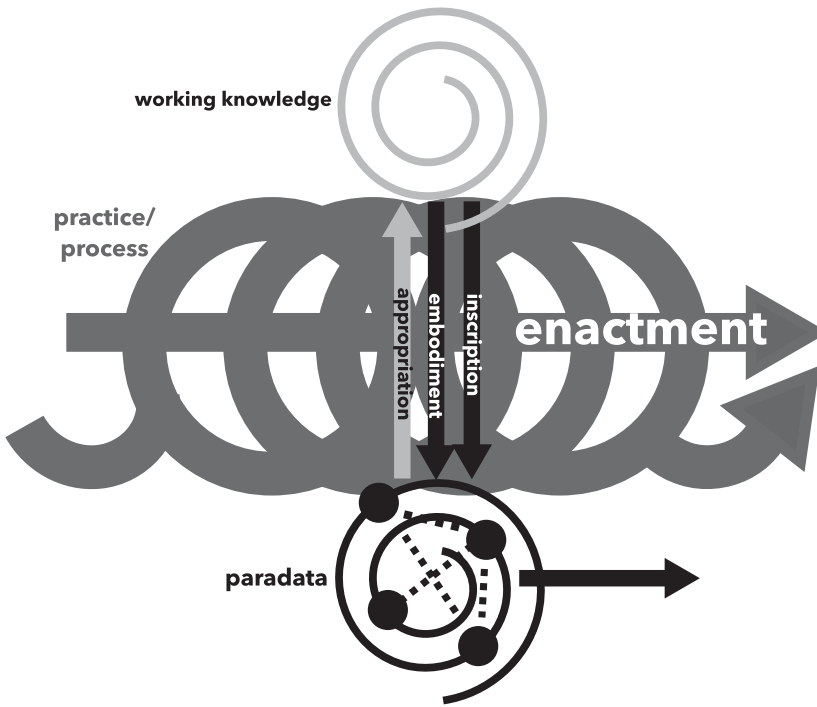

Figure 7.2 Working knowledge intersects with practices and processes.

about or even *of* but rather knowledge *in* the processes and practices themselves. The concept of working knowledge is akin to Polanyi's (1998) idea of personal knowledge – but rather than to a person, it is personal to a practice or process. Working knowledge is about how people know what they do in practice. The knowing is implicit in nature and craft-like, materially incorporated in the physicality of practices and processes. It is intrinsic, difficult to inscribe or even recognise as knowledge. Rather than being *about* a practice, working knowledge is a part of practice to an extent that distinguishing the two makes sense for analytical purposes but no more than that.

Understanding it becomes easier if we consider examples of how practices/processes and knowing are fused in practice. Researchers know how certain types of studies are conventionally expected to be done in their field. Comparably, a competent data archivist knows how data archiving is supposed to be done according to the existing guidelines, and data analysts know how they themselves and their colleagues do their work. Their knowledge incorporates both what happens in practice, including anything that can be conceptualised as steps, tasks or the workflows, but also the implicit and explicit contexts of processes and practices (Kowalczyk and Shankar, 2011), unarticulated decisions and their underpinning reasons. Essentially, working knowledge is a flow. It is about knowing rather than being.

In contrast to the working knowledge embedded in processes and practices, what we call in this volume paradata unfolds as a *category of things that can be appropriated as informative of processes and practices*. However, rather than embedded knowledge, it is a category of *process and practice information* which is to varying degrees externalised from its associated processes and practices. It is not necessarily *about* practices or processes but 'near' them (as metadata for Parry, 2023) to an extent that it has capacity to make them intelligible. In the context of this volume where the focus is on paradata describing activities pertaining to data, we have been particularly interested in things that are associated with and which can be appropriated as informative about data-related processes and practices.

This does not, however, mean that a general concept of paradata would need to apply only to data-related processes and practices when it is applied. Transparency of practices and processes is equally relevant when artificial intelligence technologies are applied to an increasing extent in decision-making, generation and summarisation of information (as per Davet et al., 2023). It has been a long-term concern in communicating the processes of making heritage visualisations (Bentkowska-Kafel et al., 2012) and a burning question with the processes underpinning and relating to generation and management of archival records (e.g., Jones and Bunn, 2024; Packalén and

Henttonen, 2024). The examples from the empirical research show how paradata can consist of physical and conceptual things including narratives and photographs in research reports (e.g., Huvila et al., 2021) or concepts used to describe observations in a dataset (Börjesson et al., 2022). It is always material to some degree through the materiality of the paradata-things and the materialities of its associated processes and practices.

There is no reason to limit potential paradata to intentional human-made documents (cf. Ferraris, 2014 and Chapter 2) even if humans perhaps have a particular appetite to develop and utilise accounts of doings – what Ferraris describes as human necessity to leave traces (Ferraris, 2014). In practical terms, paradata can be an individual narrative or visual depiction of a process or practice, or an assemblage or a sequence of such descriptions or prescriptions organised as a workflow. Many things appropriable as such are inscribable as metadata, textual descriptions or diagrams. They can be written out to provide (enough) information for an outsider to master (knowledge of) a process or practice. They can be made, captured, organised, managed and used as if paradata was a distinct entity. The malleability, durability and portability that paradata promises are major advantages. At the same time, it also comes with a potential risk of misrepresentation, incompleteness and over-simplification.

Unlike implicit and embedded working knowledge of processes and practices, when a knowledge is translated into paradata it is objectified in constellations of physical or conceptual things that span over both space and time. It forms a network (lower part in black in Figure 7.2) rather than a hierarchical structure or a monolithic entity. Rather than standing alone, paradata works together and links to a large variety of different types of information (e.g., Huvila et al., 2022). In a semiotic sense, a referent (practice or process) is encoded in a signans (paradata). The working knowledge is specified to a particular coordinate of reference and articulated either verbally or using some other means of expression (cf. Ingold, 2022).

However, rather than inherently being paradata, 'things' become paradata first when their particular kind of documentality (cf. Day, 2024; Ogden and Richards, 1930) as paradata is acknowledged. Equally, as formulated above, a thing becomes paradata when it is appropriated as such and becomes a part of a particular constructed totality (cf. Lund, 2024 on when a document becomes a document) where it functions as paradata.

As discussed in Chapters 3 and 6, the diversity of things potentially appropriable as paradata is so broad to an extent that even an absence of things, information gaps or non-information can occasionally become paradata, that is, informative of practices or processes (Huvila et al., 2023b). The idea of

paradata becoming paradata through appropriation ties in with the idea of staying within and traversing boundaries discussed in Chapter 3 in terms of technical and epistemic thresholds. Such an understanding of paradata follows essentially the line of what Huvila (2022) describes as a middle-range approach to paradata, that is, whether a thing is paradata, data or metadata depends on perspective and how it is used. One person's data can be another's paradata and vice versa.

For example, while a researcher in physics uses lab notebooks as paradata on how experiments were conducted, a science studies scholar might use them as data when doing research on physicists' research practices. Sometimes, as with notebooks (Canfield et al., 2011) and protocols (Rheinberger, 2023) used to document both the subject and process of research, making such a distinction can be impossible. Data can be paradata at the same time as it is data.

7.4 Embodied and Inscribed

In the previous section, paradata was approached as things with informational potential. The origins of such things can vary and in much the same way that Mayernik observes of metadata: Paradata is performed 'differently in different social settings and situations' (Mayernik, 2020, p. 702). When considering what kind of information paradata is, there is a distinction to make between information or documents that are made-to-be (for Hauser 2024, facta) paradata in terms of that they are *inscribed* and given as such (for Hauser 2024, data) or *embodied* and merely taken as such (for Hauser 2024, capta). This characterises also the distinction between what is discussed in Chapter 6 in terms of made-to-be paradata proper and incidental, potential paradata.

A written document or drawing of a process or practice are typical examples of inscriptions that can be appropriated as paradata. They can be formal or informal, more or less structured, but typically at least to some extent they are products planned rather than outcomes of incidental information making (cf. Chapter 3 in this volume). As Mathieu (2023) suggests, inscriptions also vary in how they link to the practices and processes. As can be sensed in the different methods and approaches discussed in Chapters 4 and 5, it is possible to see inscriptions to (cf. Mathieu 2023):

- *describe* practices or processes by conveying an account of them;
- *prescribe* by instructing or stipulating them;
- *transcribe* by rearranging them;
- *proscribe* by forbidding access to them;
- *subscribe* by adhering to them;

- *circumscribe* by restricting access to them; or
- *ascribe* by explaining them.

Elements of practices and processes can, however, also be embodied in multiple things (Baker, 2017), for instance, in datasets of results created during a research project or physical artefacts crafted by artisans. As discussed in Chapter 2, paradata is often informal and unstructured, and of secondary nature.

Pagés ([1948] 2021) introduced the notion of auto-documents to discuss the documentary function of things not purposefully created as documents (see also Buckland, 2024; Day, 2024). Auto-documents are taken-to-be (*capta*) while documents are also given-to-be (*data*) information (Hauser, 2024) or paradata when used as such. While auto-documents are akin to natural information, documentary paradata is always either embodied or encoded (as information for Bates, 2006). Ingold makes a related distinction by emphasising that deposition and palimpsests are not inscriptions (Ingold, 2022, p. 186–190). In palimpsets, 'chalk marks are not inscribed. Rather than sinking into the surface . . . the chalk is deposited' (Ingold, 2022, p. 188).

The same distinction can be made between two types of paradata: inscriptions and embodiments. These roughly correspond to what has also been termed ingredients and traces (Huvila et al., 2023). Traces and embodiments follow the stratigraphic logic of hierarchy from the present to the past, whereas ingredients and inscriptions are rhizomatic by their nature, growing from below to the surface (cf. Ingold, 2022, p. 194). This means that embodied things may inform as paradata through going back in time to the moment when a process or practice took place. In contrast, with an inscription, the journey goes forward in time from the moment when the inscription was generated to the present where it is eventually appropriatable as paradata.

There are diverse mechanisms of how paradata comes into being as inscribed or embodied. Therefore, how things turn to paradata through appropriation and how things sometimes function as paradata and sometimes not means that there is no single moment when paradata happens. This is recognised in Figures 7.1 and 7.2 in the iterative, turbulent and irregular intertwined movement of processes, practices and paradata from left to right.

Being in the making is common to both working knowledge and paradata. Understanding the differences between the flow of working knowledge and the iterative remaking of paradata can help to overcome the gap between them. Working knowledge, like habits for Ingold (2022), 'resist explication' to a degree that paradata cannot be complete.

Thinking of paradata in an objectifying sense means to try to 'specify' or 'fix coordinates of reference' (Ingold, 2022, p. 231) of a practice or process

and articulate them as being definite, essentially turning the practices and processes to aggregated models and sets of datapoints (cf. Hartley and Schjøtt, 2023). It has aspirations to be, in Vetter's (2016) terms, cosmopolitan across contexts and practices rather than strictly experiential. In contrast to assuming that paradata is according to the premises of systems-oriented knowledge management (Handzic, 2004) implicit process and practice knowledge made explicit, paradata is better understood as a verb (cf. Dervin, 1999; Latour, 2011) and being in-the-making. Such perspective recognises its incompleteness and fluidity in relation to working knowledge and the processes and practices themselves.

Being in the making means that paradata is potentially coming into being at multiple points of time. Rather than being determined only at and by the moment when a particular artefact may be paradata for someone, it is made again and again as paradata through embodiment, inscription and appropriations of things whenever they are (re)embodied, (re)inscribed or acted upon as paradata. Inscribed paradata is (re)made when it is conceived and whenever it is used.

Embodied paradata, if understood as paradata from the outset, may follow the same pattern. Sometimes things embodying working knowledge might be recognised and appropriated as paradata first at use, suggesting that it is crucial to distinguish the making of things appropriatable as paradata and paradata itself. This corresponds to how the authoring of documents and turning the documents to communicate, mediate or, for example, change practices are distinct acts (Huvila, 2019).

Approaching paradata from the perspective of being in-the-making does not mean that certain things might not be more widely recognised as paradata than others. Purposefully inscribed or recorded material (cf. Opgenhaffen, 2021) might generally be more obviously paradata than embodied practices but the dividing line is anything but clear. Earlier in this volume we have reviewed a plethora of categories of information that rather uncontroversially qualify as paradata for data creators, users and managers alike. Making a field diary or a workflow diagram to work as paradata often requires relatively little effort.

By contrast, we have seen how other artefacts like photographs of archaeologists working in the field (known as 'action shots') have potential to function as paradata, but making this eventually happen requires more work to render the artefact informative in terms of paradata, of practices and processes. The conundrum is comparable to what Sacks (1972) discusses in the context of children's stories: how something eventually can become recognisable as a possible description; how something may sound or look like a description; and how some of them unfold as recognisable descriptions. This

work is likely to require both intellectual labour to reconceptualise a thing as (potential) paradata, as well as material labour to (re)work the thing in a form that makes it enactable as paradata. Rather than being a simple act of picking up paradata and using it, it requires active information filtration work (Nielsen, 2015) and 'taking' (Huvila, 2022b) it in use.

We can consider two examples of the multiplicity of the making of paradata: writing narrative descriptions of a work process and sketching a diagram representing the same process. Writing a narrative description of a work process and making it to become (regarded as) paradata can be expected to happen often simultaneously, either when the work is being done or retrospectively after it has been completed. This is, however, only the first iteration of paradata in the making. The narrative can obviously be revised and rewritten. Such actions correspond with different degrees of (re)making the written-narrative-as-paradata. The making of that particular narrative happens again every time it is enacted, or 'taken' (Huvila, 2022a), in use as paradata to inform of the specific process or practice it is describing. This might happen once or several times in the course of time with the paradata being performed as distinct instances of information depending on when, where, for what purposes and by whom it is taken into use.

Another example discussed earlier in this volume is when a workflow diagram is drawn to prescribe a work process. It is not necessarily conceived as paradata but rather as a plan or suggestion of a possible future way of working. At this time, a thing is made whereas it is first transformed into paradata when it is appropriated in use as such. Borrowing the term of Taylor (1971), paradata is not a 'brute' form of data: its referentiality and meaning are in the eyes of its beholder.

Besides putting emphasis on the reality that paradata is never finished, treating paradata as being in-the-making also shifts attention to the complexity of the circumstances of how paradata comes into being. Following Latour (2001), rather than being constructed by a constructor, paradata can be considered as being *instaurated* in something Pickering (1995) has described as a thick of things that spans over time between the moment when paradata is made and when it is enacted (or taken) in use.

As Mathieu (2023) notes of agency, also paradata and its agency is always to different degrees prospective and retrospective in time, introspective in how it is looking into itself, respective of the norms of the domain where it operates, and suspective of potential risks. Similarly, for Latour (2011), this process of instauration works backwards from the moment it is enacted to the practice or process it is linked to, and back to the present. Instauration involves making a turn (Strathern, 2005a) from looking forward to turning backward. Paradata,

the processes or practices it describes and the actors (or actants) involved all participate in its making.

In identifying and understanding what might work as paradata, it is important – borrowing again from Pickering (1995) – to understand the mangle of practice within which the making takes place. Instead of merely asking what is paradata and who is making it, making one's way through the mangle requires raising a range of additional questions. Several scholars including Star and Ruhleder (1994) and Engeström (1990) have suggested asking *when* something is instead of *what* it is. Further, it is relevant to ask where and in what context something is turned into paradata, why and for what purpose, as well as how it is made. In the following sections, we are turning our attention to these heuristics by inquiring into the temporalities (when), enactment (how) and goals (why) of paradata.

7.4.1 In the Continuum

In the previous section we followed the movement of paradata between being made and appropriated as such. Another axis of movement in the model and the life of paradata, is the one that pertains to the temporalities of the making of different kinds of paradata. Even if the earlier literature has sometimes focused on paradata as created during and sometimes after a practice or process is taking or has taken place (e.g., in survey research, Kreuter, 2013; Schenk and Reuß, 2024), the potential temporalities of what can be termed paradata are broader. This has been touched upon already to some extent in the earlier conceptualisations of the term as discussed in Chapter 2 and goes beyond a dichotomy of past and future practices and corresponding records (cf. Duranti and Thibodeau, 2006).

As suggested in Chapters 4 and 5 of this volume, studies of paradata creation practices and information that qualifies as paradata occur on a temporal scale. Paradata can be extended to engender time before (*prospective*) and after (*retrospective*) but also during (*in situ*) a process or practice being performed (Chapters 2 and 3; see Figure 7.3). Plans, protocols, handbooks and guidelines are prospective in how they are created before the practice or process they are describing takes place. At the same time, many such artefacts tend also to be prescriptive in how they are generated in order to stipulate forthcoming practices and processes.

At the moment of creation, they emerge as a form of potential information that eventually can be appropriated as paradata of an actual process or practice when it has taken place. In addition to prospective paradata, much paradata making takes place in situ at the moment when a process or practice is happening.

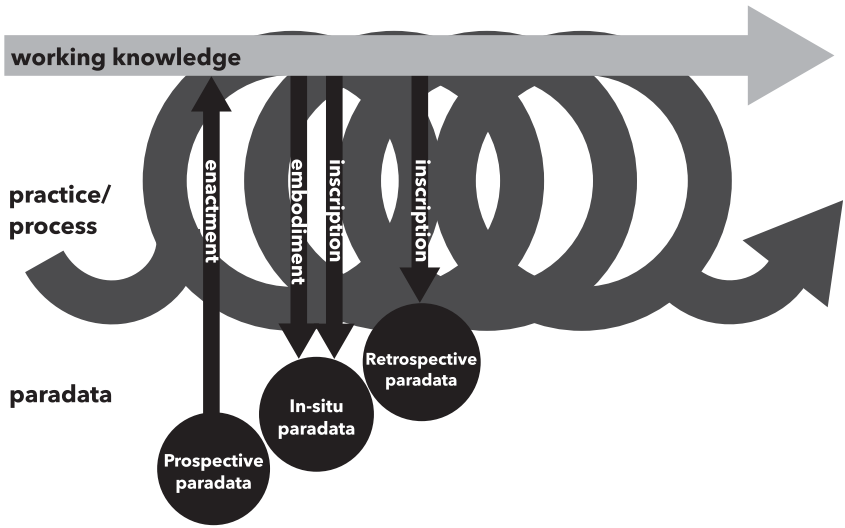

Figure 7.3 Prospective, in situ and retrospective paradata in the continuum.

As discussed earlier in this volume (Chapters 3 and 4), people write notes, keep diaries, take photos and record films – to mention only a few examples of how ongoing practices and processes can be documented. Finally, much paradata creation happens retrospectively. People write stories, draw diagrams and model previous practices and processes from memory. Sometimes paradata can be generated using various types of forensic methods (for possible methods cf. Duranti, 2009; Huvila, 2022a; Kirschenbaum et al., 2010), including metaphorically 'excavating' data, documentation and results (cf. Williams, 2011) of previous processes and practices to understand them better.

While paradata itself is not intrinsically prospective, in situ or retrospective, the moment of its making either before, during or after a particular process or practice is enacted will have implications for the resulting paradata and the limits to what it can be. This means that paradata, like a narrative (Tsoukas, 2013), is fundamentally asymmetrical. In situ observations differ from retrospective descriptions as clock time differs from narrative time (Ricoeur, 1980). An in situ account exhibits processes and practices differently than how they are remembered afterwards or conceived of beforehand.

While none of these types of accounts is necessarily better or worse than the others, the different conditions of envisioning future processes and practices (describing them while they are enacted versus attempting to recall them afterwards) mean that they differ in what they can achieve. As the

methodological exposés in Chapters 4 and 5 show, documentation in situ might allow taking note of details that are impossible to remember afterwards. Retrospective paradata generation can, for its part, benefit from a richer contextual understanding unfolding through reflection and insights gained after the process or practice has taken place. Similarly to how Tsoukas (2013) describes the relation of rules and acting upon them, it is also conceivable that the paradata that is generated in situ as a practice is enacted or afterwards is richer than paradata that is created prospectively.

The fundamental difference between paradata that precedes action and those that parallel or follow it is that at the outset, prospective paradata becomes paradata first when it is enacted as a blueprint of a practice or process that actually takes place. Prospective paradata, for example, a prescriptive plan or a predetermined workflow, turns to paradata (proper) to an extent it corresponds with the practice that actually happened. The caveat is that the account might not be completely accurate as actions may have changed as the situation developed (cf. Huvila and Sköld, 2023). It may also not reflect current practices or be an accurate guide to future ones (cf. data management plans in Mannheimer, 2022). The part that is never followed becomes what can be termed a subjunctive description (cf. subjunctive provenance for Bettivia et al., 2023 cf. how they use the term of retrospectively identifying what could have been) of processes and practices that could have happened. It is paradata of a process or practice that never was.

The obvious merit of distinguishing prospective, in situ and retrospective paradata-making is in how the three categories correspond with the temporalities of the making of informative things. This has important implications for understanding the temporal limitations of their ability to act as paradata. Our judgement of their value as more or less accurate and sometimes subjunctive accounts, however, should not be judged solely in terms of temporality of their making, although this must inevitably influence our understanding of the potential utility as paradata. Therefore, rather than assuming that it is possible to say definitely that paradata should precede, parallel or follow practices, it is more fruitful to consider what different temporalities can contribute to our understanding of them.

7.5 Enactment of Paradata

Up to this point, this chapter has delved into the mechanisms of how paradata is generated and working knowledge turns to paradata through embodiment and inscription. Our understanding shifts when we view things from a different

direction: one where paradata is *enacted* in practices and processes generating both action and working knowledge (Figure 7.2). In these enactments, paradata has agency. Comparably to how Duranti and Thibodeau (2006) distinguish instructive records from enabling ones, paradata can be *agentical* in how it provides agency to (re)enact processes and practices, learning, reuse of data and much more. It is *agentic* in how it itself has agency (cf. Hauser, 2024) to instruct, enable and influence processes and practices.

However, even if the idea of all paradata from the moment prospective paradata becomes actual paradata looks backward to a past practice, paradata functions forward (cf. Rheinberger, 2023). It should not be reified (Bachelard, 1949). Paradata cannot be used to exactly replicate the past. Paradata conveys something to someone in an unknown future from the perspective of the time when it was generated at the time when it is faced upon. It is not possible for paradata to explain or describe processes and practices from the perspective of the moment when they are put into practice in the future.

For the same reason, its quality can only be measured backwards. When it does what it is expected to do in the present in relation to a past process or practice, we can say that it is successful. This is impossible to determine in advance, neither in general nor in relation to a conceivable or unexpected future. The future-orientation means also that there are limits to what extent paradata can support what Ellingsen and Monteiro (2003) term 'rendering' both itself and practices comprehensible and actionable in an unknown future. Depending on the complexity of the processes and practices, enacting paradata might require varying amounts of repetition and experimentation (cf. Rheinberger, 2023) to succeed. Or it might fail. In addition to enacting versions of earlier practices and processes, its implementation is likely to create *collateral realities*, that is, practices and processes that are incidental and sometimes objectionable (Law, 2012).

The uncertain passage of prospective paradata from a potential account of what-is-to-be to a subjunctive or to a varying degree faithful and incomplete account of what happened in the past is illustrative not only of the fluid temporalities of paradata but also of how it informs of practices and processes in different situations. The shift of prospective paradata to a pre-scribed description shows how same artefacts can be enacted to prescribe and describe. Workflows and notebooks discussed in Chapter 4 illustrate the malleability of some approaches and genres of paradata.

The contextuality of paradata both in its making and its implementation means that it necessarily operates in and across multiple milieus that are in a constant state of flux. In terms of the Nonaka's notion of *ba* (Nonaka and Konno, 1998), a physical, virtual and mental place or space within which

information becomes knowledge, paradata needs to operate in several *ba*s at the same time to function as their key constituent.

At the same time, different perspectives on how paradata is made and enacted are particular to these places or spaces, which are more than spaces or places where paradata operates but rather arrangements of everything involved in how paradata works. Following Stengers' (2005) theorising on *ecologies of practice*, they can be linked to broader constellations of sayings, doings and beings, sometimes termed *practice architectures* (Olsson et al., 2024) or *data cultures* in data practice-related research (Huvila and Sinnamon, 2025; Oliver et al., 2023). Each of the constellations comes with their respective ecologies of knowledge (Santos, 2016) – or perhaps, more aptly, of knowing – and 'styles of thinking and doing' (Hacking, 2012).

An ecology of practice, paradata comes with a potential to form a layer – a paradata practice – that compares to what Robichaud et al., (2004) have described as a meta-conversation. It is itself layered and forms a fractal (cf. Strathern, 2005b) of nested datasets of introspection. This is also reflected in Gant and Reilly's (2017) term 'peridata' and manifested in the many connected shapes and forms of provenance data involved in archival description models (Douglas, 2017).

Just as a meta-conversation does in a discursive sense, the layers of paradata practice generate and sustain an understanding of the practices and processes it informs about. The key relations in the paradata practice are the connections between paradata makers' and paradata users' practices and processes (cf. Robichaud et al., 2004).

Ideally, the continuum of practices and processes, paradata and paradata practice would be self-reproductive, autopoietic (cf. Maturana and Varela, 1980), without a need for additional explicit actions to sustain it. In practice, it is often far from being the case, as the frequent experience of a lack of process knowledge demonstrates. Rather than autopoietic, the continuum is sympoietic (as for Krippendorff, 2023), that is, dependent on its constituents, including paradata and its related practices and processes. It should also stay as an 'ecology of partial connections' (Stengers, 2018) in the sense that it remains open to critique and helps data making and use to change whenever data practices change.

As Hall (2007 [1973]) suggests of media production and consumption, paradata making and use also need to be studied in their own right. Knitting together a practice or process with paradata, whether it is a diagrammatic model or a narrative, requires curation in terms of discerning relevant readings of paradata and caring for it through repair and maintenance work (cf. Pink et al., 2018). Chapter 6 uncovers some of the intricacies of what keeping

functioning paradata requires and how the management of paradata is intricately intertwined with how it is made to happen before, during and after its production.

7.6 Outcomes of Paradata

A relevant follow-up question to how paradata informs practices and processes is: what is paradata capable of achieving? The literature (e.g., Börjesson et al., 2022; Huvila and Sinnamon, 2022) and earlier chapters in this volume have described diverse practical desired outcomes of describing practices and processes. They include enabling and facilitating data-sharing, data reuse, reproducibility of research, verification of the outcomes and the quality of a particular process, and understanding a particular dataset and earlier interpretations based on the data. On a theoretical note, the question of paradata outcomes has somewhat different dimensions. Rather than being the outcome of what can be achieved *with* paradata, the fundamental question is what paradata itself might be capable of achieving.

With respect to what paradata is good for, independent of its context and domain, the common denominator is greater transparency (Sköld et al., 2022). What exactly is meant by transparency varies according to the ambitions of how much transparency is considered desirable or achievable. The introduction chapter of this book discusses some of the cultural and political underpinnings of the contemporary sense of urgency for particular kinds of transparencies, and how paradata stands out as a particularly opportune response to such aspirations.

The sometimes explicit but often unspoken ideal underpinning paradata discourse is an expectation of complete transparency. Entertaining the idea of what complete transparency and absolute opacity might entail in practice has obvious value as a thought experiment. It is also apparent that it is, as Bowker (2005) noted of raw data, both an oxymoron and a bad idea.

On one level, it is hardly desirable: both transparency and paradata are political and have a tendency to be unevenly beneficial to those it concerns. Actors with a lot of resources and uncontrolled power tend to prevail, and vulnerable and democratic systems lose (Adams, 2020). Even if limiting transparency is equally political and difficult, and has both positive and negative consequences, it is worth trying to imagine what Bates et al. (2023) term as meaningful and socially meaningful rather than total transparencies.

This is in part because total transparency is also hardly achievable in practice. Neither the continuum of working knowledge nor the circuit of generating and

appropriating paradata can capture everything without any loss. The knowledge of practices is at best good enough and even as such, it is difficult to say how good it is and on what premises (cf. Huvila, 2012). Prospective forms of paradata might not correspond well with what eventually happens in the future, in situ generation of paradata cannot possibly capture everything, and retrospective paradata is subject both to amnesia and a false impression comparable to Bourdieu's (1986) biographical illusion, that is, representing the past as a deceptively coherent line of events leading to the present.

A parallel question to what type or level of transparency paradata might contribute is how paradata contributes to it. In this volume, we have rejected the objectivist aspiration of being able to turn implicit working knowledge to explicit paradata and back. Rather than achieving transparency by its presence (being), theorising paradata on a continuum (embodied, inscribed and enacted) suggests that it accomplishes transparency through a reciprocal enactment in the thick of things, incorporating not only paradata itself but also its makers and users.

As a form of prospective, contemporary or retrospective historical discourse (White, 1980) informative of processes and practices, paradata also both narrates by unfolding processes and practices, and narrativises by imposing on them a particular form either before, during, or afterwards. It has agency and capacity both to describe and create processes and practices. It does this through reification (Bachelard, 1949) or sedimentation (Husserl, 1989) of working knowledge on processes and practices (cf. Rheinberger, 2023).

As for design (cf. J. Bardzell and S. Bardzell, 2013; Dunne and Raby, 2001), forming a *lens*, paradata enables and facilitates critique and critical understanding of paradata generation practices as well as of the practices and processes themselves. For example, a description of a data collection episode can help to assess its validity, understand bias, and assess to what extent the generated data might be usable for addressing other research questions.

When operating as a form of data rather than a lens, paradata becomes a different type of a device: either affirmative of a particular process or practice, or a problem-solving one to address a certain predicament or obstacle. In this latter sense, as a form of data rather than a lens, the same description of the data collection episode operates as an *explanation*, either causal (description of a practice and/or process and evidence to support it) or teleological (what is the paradata's purpose). It does not *help* to understand as it does as a lens but it is expected to explain to an extent that conveys an understanding.

In practice, paradata can be mobilised as information to solve a problem but it can also serve an orientational purpose as a kind of domain information for a specific epistemic culture (cf. Korkeamäki et al., 2024). As an explanation (cf.

Berry, 2023), paradata faces a problem of how to represent practices and processes in such a way that make the explanations meaningful in relation to particular goals of both their producers and users.

As means to what Garfinkel has described as social groups' making of their activities 'visibly-rational-and-reportable-for-all-practical-purposes' (Garfinkel, 1967, p. vii) for themselves, paradata is meaningful in a very different sense than when its goal is to mitigate anxieties or to convey a comprehensive understanding of practices or processes. Explaining activities for insiders requires a different set of cues than explaining them to outsiders. If the goal of paradata is to mitigate external anxieties, explanation does not necessarily need to be complex, as providing a comprehensive understanding of a practice or process demands a lot of both the explanations and the sophistication of its intended 'user' (cf. Berry, 2023).

Further, while the usefulness of paradata as an explanation should not be underestimated, it might be an overtly unambitious aim to suggest that it suffices to make paradata helpful. It is reasonable to argue, as Berry (2023) demands of digital infrastructures, that the goal of comprehensive paradata should not be to function as mere explanations but rather to make processes and practices understandable. This is also what the paradata literature tends to insinuate. For example, heritage visualisation literature refers to 'intellectual transparency' as 'the provision of information, presented in any medium or format, to allow users to understand the nature and scope of a "knowledge claim" made by a computer-based visualization outcome.' (Bentkowska-Kafel et al., 2012, p. 262; also Forte and Pescarin, 2012). Paradata is expected to make it possible for users to understand instead of merely informing about the intellectual underpinnings of what is being described.

The multiplicity of the possible uses of paradata is, however, obviously also a part of the conundrum that makes it a wicked problem. Paradata might appear as a solvable problem as long as we only focus on either its makers or users but turns to a wicked one as soon as both communities' interests are brought in. A major challenge is to determine what is the 'information' (cf. Bentkowska-Kafel et al., 2012, p. 262) that 'allow[s] users to understand the nature and scope of a "knowledge claim"'.

While the general goals around what paradata is expected to achieve have a lot of similarities across domains, paradata research has made it apparent that the views on what exactly is needed to enable, reuse or verify the quality of outcomes differs between data creators and reusers and between different schools of data use. Research on data practices shows that from data creators' and users' perspectives, the aspects of data-related processes and practices that require explanation differ from each other (Lian et al., 2023; Thoegersen, 2018).

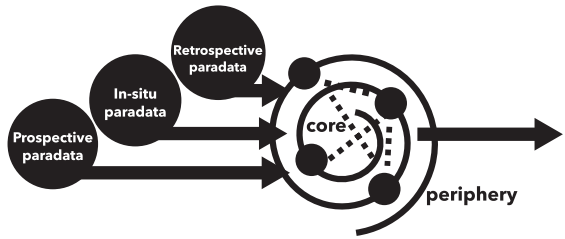

Figure 7.4 Paradata as a network.

Similarly, the requirements of paradata are different for a data user conducting statistical data analysis than for one engaged in ethnographic writing. In this respect, the key question is what kind of paradata is most useful and when.

However, as paradata does not merely exist but also does things, it can be meaningful even when it (invariably) fails to incorporate a complete understanding of a practice or process. There are other possibilities than everything or nothing. The network or meshwork-like nature of paradata (discussed earlier and illustrated in the Figures 7.1 and 7.4) means that paradata can indeed serve more than one purpose. The different things acting as such can open up parallel vistas to practices and processes and work in parallel tracks towards multiple goals instead of only one. Apart from simplifying and making something quantifiable, paradata can help to revive uncertainty and complexity.

However, while paradata may have no single goal, thinking of paradata as an explanation underlines the primacy of explaining and explanations rather than ensuring or even aiming at transparency, explaining or understanding. Explanations might not be able to make visible what Ingold (2022, p. 262) argues formal articulations fail to do, that is, to expose the *work* in practice or process, but they might still explain where and what the work is about.

Similarly crucial to making paradata meaningful is that it is clear what practices or processes paradata describes. The long temporal span from, multimodality and broad contextual scope of paradata makes it difficult. In this respect, more important than worrying about the changes of vocabulary and connotations of individual descriptive terms is that the explanations in paradata remain clear in what they explain.

7.7 In the Borderlands: Implications for Paradata Practice

The diversity of the various things that can be appropriated as informative of processes and practices and the malleability of the ways they work means that

the idea of paradata envisioned in this chapter and volume lends itself to multiple renditions. Even if the naive idea that paradata can act as a complete surrogate of a practice or process must be rejected, the concept of paradata still lends itself to be used in widely different epistemes. In a meticulously defined and parameterised context, a structured set of formally standardised paradata can operate as a script that holds together a perfect 'data supply chain' (cf. Spanaki et al., 2018). In an open world, practice and process information has a lot of friction, remains unfenced and refuses attempts to formalise it. Depending on the setting, paradata might take different forms and is interpreted differently. As a whole, it forms a network, or meshwork of practice. This meshwork operates from the outset of what Ettema and Glasser (1998) describe in the context of journalistic practices as *factual coherence*. Knowledge of a practice or process is formed through triangulating individual sources of information and looking at how they fit together and form a coherent whole, rather than relying on any one source.

Even if the meshwork of paradata lacks a definite form, it is radial and centred around an intersection of prototypical core elements in a given context and malleable non-prototypical periphery of elements at different conceptual distances from its nucleus (Figure 7.5). Core and peripheral elements and their value are interpreted according to the setting.

In laboratory research the core elements might consist of the documentation of instrumentation and formal steps of analyses whereas contextual information on those involved in the process and the laboratory itself could provide more peripheral but still valuable complements to the meshwork of paradata. In archaeological field documentation (Huvila et al., 2021), a textual, diagrammatic and photographic description of work process could form the core whereas citations to the literature might occupy a more peripheral position. In both cases, the formal core documentation might be considered as more authoritative as an account of laboratory or field practices for laboratory research and field archaeology. However, outside of such, to a varying extent

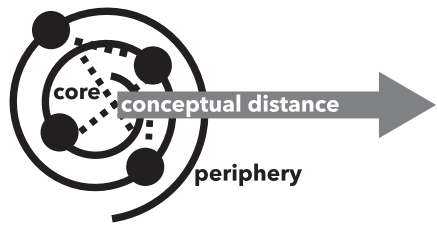

Figure 7.5 Core paradata forming a meshwork with peripheral elements at different conceptual distances from the core.

specific settings, the core and nucleus might shift radically. For a scholar of science and scholarship, the peripheral cues might form the core for understanding research practice and the formal accounts be much less interesting.

At the same time as being radial and centred, paradata exists somewhere between its makers and users, the situations where paradata emerges and is exploited, and ultimately between the practices and processes it refers to and within which it is used. Like the archaeological record according to Barrett (1988), paradata surfaces as a field of fragmentary traces of social practices and their relationships rather than as a clean narrative. In this space and state of in-betweenness, paradata engages in the boundary work of bringing people, practices and processes together and pushing them apart. Good paradata works as what Star (1988; 2010) terms as a boundary object.

Rather than forming an 'actuarial record' of a practice or process, paradata could do better acting – as Dallas has proposed of archaeological 3D visualisations building on Garfinkel (1967) – as an 'epistemic contract' between neighbouring contexts and communities. It needs to remain liminal and worked to stay as such, somewhere between tacit and explicable or specified (cf. Ingold, 2022, p. 231–232) to be effective across the contexts and communities where it can play a meaningful role.

Paradata is also characterised what Currie (2023) describes in his study of historiography as the difference between narratives and chronology. Currie (2023) notes that two parties might agree on chronology, that is, the sequence of certain historical events but at the same time find disagreement between narratives of what they mean and what is their significance and relation to other historical events. While the typical primary concern of knowledge organisation tends to be on achieving what might be compared to getting the chronology straight, paradata comes with an opportunity to focus on what can be likened to narratives, that is, providing context to an extent that makes it informative of complex practices and processes. The comparison to chronologies and narratives and Currie's (2023) observation that it is possible to find agreement between historical chronologies of events but simultaneously disagreement between historical narratives of them can be similarly transposed to elaborate on the paradoxical sufficiency and inadequacy of parallel forms of paradata in specific situations. It compares to how different forms of paradata can at the same time be adequate and in agreement on the level of technical details but at the same time inadequate in conveying a practice or process in its full complexity for different individuals or communities.

Thinking of paradata from the outset of the reference model residing in the borderlands helps also to explicate its impact in practice across the information field. Paradata has apparent value for advancing the *ideals of the open*

movement including open science, open data and knowledge. However, it is equally obvious that the transparency and openness paradata generates is not unconditional and absolute in how it emanates from documentation or availability of information. Thinking back how paradata acts (cf. Mathieu, 2023), to mediate across the epistemic borderlines of practice architectures or data cultures in the time-space of practices and processes, paradata needs to look backwards and forwards, it needs be introspective and extrovert – or perhaps 'extrospective' – of itself and its context, respective of the norms of the domains where it operates, and conscious of potential risks. Accordingly, when working with paradata, it is vital to be serious about its potential desirable and undesirable consequences.

Paradata is a strong candidate to be one of the central keywords of the contemporary datafied experience. Data can hardly be called open if little or nothing is known about the processes and practices through which it came into being and through which it has been processed and used. The same also applies to 'data quality', which can be imagined to exist if the underpinning data remain concealed. The presence of paradata in the broad meaning of how the concept is used in this volume also forms a requisite to making data findable, accessible, interoperable and reusable (FAIR as for Wilkinson et al, 2016). This can only happen if it can accommodate the diversity of epistemes where data originates and is used. The crux of FAIRness and the necessary paradata to support is to acknowledge that FAIRness is not universal. If a dataset is FAIR for someone, it is probably MEAN, or miscellaneous, exceptional, arbitrary, and nonconformist, for many others (Huvila, 2017).

While paradata is indispensable both as a lens for dealing in practice with large data quantities, or big data, its most significant implication is to make and keep visible how big data really is an oxymoron. From this perspective, it has the potential to affect transformative change in *research practice* across disciplines.

For researchers using existing datasets, paradata opens up opportunities to better understand the processes and practices that data is a part of. It is not only a question of provenance (i.e. where data comes from) but also of a comprehensive understanding of the practices and processes as an element of what the data is. For researchers creating data for preservation, a major challenge and opportunity is to embrace the different means of making and keeping paradata to facilitate future enactment of practices.

Rather than thinking about paradata as an auxiliary attachment, paradata should rather be seen as a parallel to the results of a study, and a process of conveying knowledge and understanding rather than a mere output. It should not be generated and preserved only to assure others of the validity and

relevance of the results but also to provide ingredients for generating a biography of the process itself for later use.

Being in the borderlands has both its advantages and its risks. A major advantage comes with being in the middle of complex circumstances. By being there, paradata both simplifies and complicates understanding practices and processes at the same time. Like all data in general (Mejias, 2023), it facilitates control by generating differences that are exploitable in both traversing and erecting boundaries.

However, working with paradata in the borderlands also leads to higher transaction costs (Williamson, 1981) compared to having the necessary working knowledge. Generating, capturing and keeping but also putting paradata to use requires time, effort and resources. This is especially true when working simultaneously with formal and informal paradata that operate on fundamentally different terms. One question is how to design paradata while keeping transaction costs acceptable. In this respect, paradata has implications for the *design of knowledge organisation systems*. Automation and technological tools can mitigate some of the costs, for example, by providing and helping to maintain an overview of available information and keeping track of what a user has already consulted.

In contrast, information technologies generally struggle to support complex sociomaterial practices (Ellingsen and Monteiro, 2003). Formalisation can make paradata more manageable but it also makes it fixed, 'sanitized and re-represented' (Batist, 2023, p. 8) rendering invisible its instability and precariousness (Rheinberger, 2023) as was discussed in the Chapter 6. Formal paradata ascribes to the contradiction raised by Ingold that fixed and formally articulated working knowledge is that which is silent, static and inexplicable – in effect tacit – whereas working knowledge is 'turbulent and sometimes noisy' (Ingold, 2022, p. 232). While fixity of formal and structured paradata can help to overcome certain barriers to knowing, the translation from fluid to structured silences makes inexplicable many aspects of the practices and processes it documents.

Stories and other informal forms of paradata have different qualities. As Ingold argues, they 'allow experts to tell what they know *without* specifying it' (Ingold, 2022, p. 233) (emphasis in the original). Further on at the opposite end of the scale from structured standardised descriptions, like embodied craft, also processes and practices are themselves another 'a way of telling' (Ingold, 2022, p. 234) in how they convey and sustain working knowledge as a part of their existence.

As a result of the respective advantages and shortcomings of individual varieties of paradata, one form of paradata is seldom enough. Even seemingly

complete 'recipes' and scripted workflows (Gupta, 2020) miss a lot of details. Both Duranti and Thibodeau (2006) and Ingold (2022) discuss the example of a music score and its relation to practice to illustrate the gap in between. Through a paradata lens, the score does rather uncontroversially qualify as paradata but what it does is that it 'specifies the elements of a completed work and shows how they articulate. Moreover it is silent' (Ingold, 2022, p. 250). A score is not enough to reproduce a performance. Both complementary inscriptions, embodiments and working knowledge is needed before a score can be enacted to music. They all are mobile in their particular ways.

Ingold contrasts 'computational' algorithms to 'ambulatory' practices (Ingold, 2022, p. 283) but it is easy to conceive a plethora of additional means in which the different forms of paradata discussed throughout the pages of this volume are mobile in their own terms. This poses an obvious challenge for *information systems design and information and knowledge management*. If all paradata is structured and standardised, much of the critical information is bound to disappear.

Rather than suggesting that the aim of thinking and working with paradata is to replace living practices and processes with embodiments and inscriptions, its aim should perhaps be seen as appreciating them as they are and keeping them apart but in correspondence with each other allowing us, as Feinberg suggests, to 'live with referential ambiguities' (Feinberg, 2022, p. 91). As she continues, data is too often thought of in terms of surrogates rather than 'networks, collectives, or other configurations of thingness' (Feinberg, 2022, p. 92). Information systems designed for paradata should not try to suppress one with the other but use formal means to provide access to and keep track of the miscellaneity of embodiments and inscriptions of practices and processes bringing together the human and systems-oriented approaches to information and knowledge management (cf. Huvila et al., 2023). Instead of transformative boundary crossings where practices and processes change shape, paradata is more likely to thrive when left meandering in the borderlands between the practices and processes inscribed or embodied and enacted. In a very fundamental sense, paradata should remain as much a question as an answer.

7.8 Conclusions

The purpose of this chapter has been to theorise and stage a reference model for paradata and how it is practised. Unlike in much of the earlier literature, the model presented here puts forward an understanding of paradata as a parallel mode of being where the practices and processes inscribe and embody. It goes

side by side with working knowledge in practices and processes that themselves engender paradata through embodiment and acts of inscription. Paradata turns back to working knowledge through appropriation and through enactment to practices and processes. Paradata is perpetually in the making on a continuum, and actualises as instances only through temporal intersections with the practices and processes it causes to converge. In parallel to being processual, consisting of interlinked types of information connected to each other, paradata forms a network-like meshwork where the whole is larger than its individual constituents.

In contrast to how paradata is typically defined as a means to provide transparency, this chapter has problematised that idea, suggesting that it might instead be more fruitful to approach it in terms of explanations, acknowledging at the same time its spanning over the divide between practices and processes and knowledge thereof.

Rather than taking paradata as a script or comprehensive account of a practice or process, it can be more fruitful to leave it in the borderlands between doings and descriptions. Being there it can retain its malleability and continue to refer in both conceptual and practical sense to a much broader range of things (appropriable informative of practices and processes) than when essentialised as a specific category of information. Rather than using the concept of paradata to claim that knowing about practices and processes is a solvable, straightforward question, paradata does better by foregrounding the opposite: the unbearable complexity of knowing in practice.

References

Adams Rachel (2020). *Transparency*. London: Routledge.

Bachelard Gaston (1949). *Le Rationalisme Appliqué*. Paris: Presses Universitaires de France.

Baker Malcolm (2017). Epilogue: Making and knowing, then and now. In Smith Pamela H., Meyers Amy R. W. and Cook Harold John (eds.), *Ways of Making and Knowing: The Material Culture of Empirical Knowledge*. Ann Arbor: University of Michigan Press, 405–413.

Bardzell Jeffrey and Bardzell Shaowen (2013). What is 'critical' about critical design? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI 13. New York: ACM, 3297–3306.

Barrett John C. (1988). Fields of discourse: Reconstituting a social archaeology. *Critique of Anthropology* 7(3), 5–16.

Bates Jo et al. (2023). Socially meaningful transparency in data-based systems: Reflections and proposals from practice. *Journal of Documentation* 80(1), 54–72.

Bates Marcia J. (2006). Fundamental forms of information. *Journal of the American Society for Information Science and Technology* 57(8), 1033–1045.

Batist Zachary (2023). *Archaeological Data Work as Continuous and Collaborative Practice*. PhD thesis. Toronto: University of Toronto.

Bentkowska-Kafel Anna, Denard Hugh and Baker Drew (eds.), (2012). *Paradata and Transparency in Virtual Heritage*. Farnham: Ashgate.

Berry David M. (2023). The explainability turn. *Digital Humanities Quarterly* 017(2). ISSN: 1938-4122.

Bettivia Rhiannon, Cheng Yi-Yun and Gryk Michael (2023). What does provenance LACK: How retrospective and prospective met the subjunctive. In Sserwanga Isaac et al. (eds.), *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity*. Vol. 13972. Cham: Springer Nature Switzerland, 74–82.

Börjesson Lisa, Huvila Isto and Sköld Olle (2022). Information needs on research data creation. *Information Research* 27. Special Issue, isic2208.

Börjesson Lisa, Sköld Olle et al. (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), 1–18.

Bourdieu Pierre (1986). L'illusion biographique. *Actes de la Recherche en Sciences Sociales* 62(1), 69–72.

Bowker Geoffrey C. (2005). *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.

Buckland Michael (2024). Revisiting Robert Pagès: Documents and culture. *Proceedings from the Document Academy* 10(2).

Canfield Michael R. et al. (2011). *Field Notes on Science and Nature*. Cambridge, MA: Harvard University Press.

Currie A. (2023). Narratives, events & monotremes: The philosophy of history in practice. *Journal of the Philosophy of History*, 17(2), 265–287.

Davet Jeremy, Hamidzadeh Babak and Franks Patricia (2023). Archivist in the machine: Paradata for AI-based automation in the archives. *Archival Science* 23(2), 275–295.

Day Ronald (2024). Powerful particulars as 'autodocuments' in documentality. *Proceedings from the Document Academy* 10(2).

Dervin Brenda (1999). On studying information seeking methodologically: The implications of connecting metatheory to method. *Information Processing and Management* 35(6), 727–750.

Douglas J. (2017). Origins and beyond: The ongoing evolution of archival ideas about provenance. In Macneil H. and Eastwood T. (eds.), *Currents of Archival Thinking*. Libraries Unlimited, 25–52.

Dunne Anthony and Raby Fiona (2001). *Design Noir: The Secret Life of Electronic Objects*. Basel, Switzerland : London: Birkhuser ; August Media Ltd.

Duranti Luciana (2009). From digital diplomatics to digital records forensics. *Archivaria* 68.Fall, 39–66.

Duranti Luciana and Thibodeau Kenneth (2006). The concept of record in interactive, experiential and dynamic environments: The view of InterPARES. *Archival Science* 6(1), 13–68.

Ellingsen G. and Monteiro E. (2003). Mechanisms for producing a working knowledge: Enacting, orchestrating and organizing. *Information and Organization* 13(3), 203–229.

Engeström Yrjö (1990). When is a tool? In *Learning, Working and Imagining, Twelve Studies in Activity Theory*. Helsinki: Orienta-Konsultit Oy, 171–195.

Ettema J. S. and Glasser T. L. (1998). *Custodians of Conscience: Investigative Journalism and Public Virtue*. Columbia University Press.

Feinberg Melanie (2022). *Everyday Adventures with Unruly Data*. Cambridge, MA: MIT Press.

Ferraris Maurizio (2014). *Documentalità: Perché è necessario lasciar tracce*. Gius: Laterza & Figli Spa.

Forte Maurizio and Pescarin Sofia (2012). Behaviours, interactions and affordances in virtual archaeology. In Bentkowska-Kafel Anna, Denard Hugh and Baker Drew. (eds.), *Paradata and Transparency in Virtual Heritage*. Farnham: Ashgate, 189–201.

Gant Stefan and Reilly Paul (2017). Different expressions of the same mode: A recent dialogue between archaeological and contemporary drawing practices. *Journal of Visual Art Practice* 17(1), 100–120.

Garfinkel Harold (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.

Golder Scott and Huberman Bernardo A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science* 32(2), 198–208.

Gupta Neha (2020). Preparing archaeological data for spatial analysis. In Gillings Mark, Hacıgüzeller Piraye and Lock Gary. *Archaeological Spatial Analysis: A Methodological Guide*. 1st ed. Routledge, 17–40.

Hacking Ian (2012). 'Language, truth and reason' 30 years later. *Studies in History and Philosophy of Science Part A. Part Special Issue: Styles of Thinking* 43(4), 599–609.

Hall S. (2007 [1973]). Encoding and decoding in the television discourse. In Gray A., Campbell J., Erickson M., Hanson S. and Wood H. (eds.), *CCCS Selected Working Papers*. Routledge, 402–414.

Handzic Meliha (2004). *Knowledge Management: Through the Technology Glass*. Singapore: World Scientific.

Hartley Jannie Møller and Schjøtt Anna (2023). Imagining publics through emerging technologies. In Møller Hartley Jannie, Sørensen Jannick Kirk and Mathieu David (eds.), *DataPublics: The Construction of Publics in Datafied Democracies*. Bristol: Bristol University Press, 99–120.

Hauser Elliott (2024). Making-to-be: Documents, facta, and material-discursive agency. *Proceedings from the Document Academy* 10(2).

Husserl Edmund (1989). The origin of geometry. In Derrida Jacques (ed.), *Edmund Husserl's Origin of Geometry, an Introduction*. Lincoln: University of Nebraska Press, 157–180.

Huvila Isto (2012). *Information services and digital literacy: In Search of the Boundaries of Knowing*. Oxford: Chandos.

Huvila Isto (2017). Being FAIR When Archaeological Information Is MEAN: Miscellaneous, Exceptional, Arbitrary, Nonconformist. *Presentation at the Centre for Digital Heritage Conference 2017, Leiden June 14–16, 2017*.

Huvila Isto (2019). Authoring social reality with documents: From authorship of documents and documentary boundary objects to practical authorship. *Journal of Documentation* 75(1), 44–61.

Huvila Isto (2022a). Improving the usefulness of research data with better paradata. *Open Information Science* 6(1), 28–48.

Huvila Isto (2022b). Making and taking information. *JASIST* 73(4), 528–541.

Huvila I., Andersson L., Fulton C., Haider J. and Harviainen J. T. (2023b). Managing information gaps and non-information. *Proceedings of the Association for Information Science and Technology*, 60(1), 793–798. https://doi.org/10.1002/pra2.863

Huvila I., Börjesson L. and Sköld O. (2022). Archaeological information-making activities according to field reports. *Library & Information Science Research*, *44*(3), 101171. https://doi.org/10.1016/j.lisr.2022.101171

Huvila Isto and Sinnamon Luanne (2022). Sharing research design, methods and process information in and out of academia. *Proceedings of the Association for Information Science and Technology* 59(1), 132–144.

Huvila Isto and Sinnamon Luanne (2024). When data sharing is an answer and when (often) it is not: Acknowledging data-driven, non-data, and data-decentered cultures. *Journal of the Association for Information Science and Technology*, 75(13), 1515–1530.https://doi.org/10.1002/asi.24957

Huvila Isto and Sköld Olle (2023). A fieldwork manual as a regulatory device: Instructing, prescribing and describing documentation work. *Journal of Information Science*, p. 01655515231203506.

Huvila Isto, Sköld Olle and Andersson Lisa (2023a). Knowing-in-practice, its traces and ingredients. In Cozza Michela and Gherardi Silvia (eds.), *The Posthumanist Epistemology of Practice Theory: Re-imagining Method in Organization Studies and Beyond*. Cham: Palgrave MacMillan, 37–69.

Huvila Isto, Sköld Olle and Börjesson Lisa (2021). Documenting information making in archaeological field reports. *Journal of Documentation* 77(5), 1107–1127.

Ingold Tim (2022). *Imagining for Real: Essays on Creation, Attention and Correspondence*. Abingdon: Routledge.

Jones, Kevin and Bunn Jenny (2024). Mapping accessions to repositories data: A case study in paradata. In Huvila Isto, Andersson Lisa and Sköld Olle (eds.), *Perspectives on Paradata: Research and Practice of Documenting Data Processes. Knowledge Management and Organizational Learning*. Cham: Springer.

Kirschenbaum Matthew G., Ovenden Richard and Redwine Gabriela (2010). *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Tech. rep. Washington, DC: CLIR.

Korkeamäki Laura, Keskustalo Heikki and Kumpulainen Sanna (2024). Types of Domain and Task-solving Information in Media Scholars' Data Interaction. *Journal of the Association for Information Science and Technology*, asi.24863.

Kowalczyk Stacy and Shankar Kalpana (2011). Data sharing in the sciences. *ARIST* 45(1), 247–294.

Kreuter Frauke (2013). Improving Surveys with Paradata: Introduction. In Kreuter Frauke (ed.), *Improving Surveys with Paradata Analytic Uses of Process Information*. Hoboken, NJ: Wiley, 1–9.

Krippendorff Klaus (2023). A critical cybernetics. *Constructivist Foundations* 19(1), 82–93.

Latour Bruno (1987). *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.
Latour Bruno (2009). *The Making of Law: An Ethnography of the Conseil d'Etat*. Oxford: Polity Press.
Latour Bruno (2011). Reflections on Etienne Souriau's Les différents modes d'existence. In Harman Graham, Bryant Levi and Srnicek Nick (eds.), *The Speculative Turn: Continental Materialism and Realism*. Melbourne: re.press, 304–333.
Law John (2012). Collateral realities. In Domínguez Rubio Fernando and Baert Patrick (eds.), *The Politics of Knowledge*. London: Routledge, 156–178.
Lian Zhiying, Oliver Gillian and Chen Yi (2023). Genres of online COVID-19 information and government information culture: A comparative case study. *Library & Information Science Research* 45(4), 101263.
Lund N. W. (2024). *Introduction to Documentation Studies*. Facet.
Mannheimer Sara (2022). Data Curation for Qualitative Data Reuse and Big Social Research. PhD Thesis. Humboldt-Universität zu Berlin.
Mathieu David (2023). Deconstructing the notion of algorithmic control over datapublics. In Møller Hartley Jannie, Sørensen Jannick Kirk and Mathieu David (eds.), *DataPublics: The Construction of Publics in Datafied Democracies*. Bristol: Bristol University Press, 27–48.
Maturana Humberto R. and Varela Francisco J. (1980). *Autopoiesis and Cognition : The Realization of the Living*. Dordrecht: Reidel.
Mayernik Matthew S. (2020). Metadata. *Knowledge Organization* 47(8), 696–713.
Mejias Ulises A. (2023). Notes on the Historiography of Data Colonialism. In Filimowicz Michael (ed.), *Decolonizing Data: Algorithms and Society*. London: Routledge, 1–14.
Nielsen Karen Dam (2015). Involving patients with e-health: The dialogic dynamics of information filtration work. *Science & Technology Studies* 28(2), 29–52.
Nonaka Ikujiro and Konno Noboru (1998). The concept of 'ba': Building a foundation for knowledge creation. *California Management Review* 40(3), 40–54.
Ogden C. K. and Richards I. A. (1930). *The Meaning Of Meaning*. London: Kegan Paul, Trench, Trubner.
Oliver Gillian et al. (2023). Understanding data culture/s: Influences, activities, and initiatives. *Annual Review of Information Science and Technology* 75(3), 201–214.
Olsson Michael, Sköld Olle and Andersson Lisa (2024). Layers Upon Layers: Data Reuse Challenges in Archaeological Contexts. In *Proceedings of the ISIC 2024*.
Opgenhaffen Loes (2021). Visualizing archaeologists: A reflexive history of visualization practice in archaeology. *Open Archaeology* 7(1), 353–377.
Packalén Saara and Henttonen Pekka (2024). Adding paradata about records processes via information control plans. In Huvila Isto, Andersson Lisa and Sköld Olle (eds.), *Perspectives on Paradata: Research and Practice of Documenting Data Processes. Knowledge Management and Organizational Learning*. Cham: Springer.
Pagès Robert ([1948] 2021). Transformations documentaires et milieu culturel. *Proceedings from the Document Academy* 8(1), Article 3.
Parry Kyle (2023). Metadata is not data about data. In Filimowicz Michael (ed.), *Decolonizing Data: Algorithms and Society*. London: Routledge, 15–33.

Pickering Andrew (1995). *The Mangle of Practice: Time, Agency, and Science*. Chicago: University of Chicago Press.

Pink Sarah et al. (2018). Broken data: Conceptualising data in an emerging world. *Big Data & Society* 5(1), p. 2053951717753228.

Polanyi Michael (1998). *Personal Knowledge*. London: Routledge.

Rheinberger Hans-Jörg (2023). *Split and Splice: A Phenomenology of Experimentation*. Chicago: University of Chicago Press.

Ricoeur Paul (1980). Narrative time. *Critical Inquiry* 7(1), 169–190.

Robichaud Daniel, Giroux Hélène and Taylor James R. (2004). The metaconversation: The recursive property of language as a key to organizing. *Academy of Management Review* 29(4), 617–634.

Sacks Harvey (1972). On the analyzability of stories by children. In Gumperz John J. and Hymes Dell (eds.), *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Rinehart & Winston, 325–345.

Santos Boaventura de Sousa (2016). *Epistemologies of the South: Justice against Epistemicide*. London: Routledge.

Schenk Patrick Oliver and Reuß Simone (2024). Paradata in surveys. In Huvila Isto, Andersson Lisa and Sköld Olle (eds.), *Perspectives on Paradata: Research and Practice of Documenting Data Processes*. Knowledge Management and Organizational Learning. Cham: Springer.

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

Spanaki Konstantina et al. (2018). Data supply chain (DSC): Research synthesis and future directions. *International Journal of Production Research* 56(3), 4447–4466.

Star Susan Leigh (1988). The Structure of Ill-Structured Solutions: Heterogeneous Problem-Solving, Boundary Objects and Distributed Artificial *Intelligence. In Proceedings of the 8th AAAI Workshop on Distributed Artificial Intelligence, Technical Report, Department of Computer Science, University of Southern California*. Los Angeles, CA, 37–54.

Star Susan Leigh (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology & Human Values* 35(5), 601–617.

Star Susan Leigh and Ruhleder Karen (1994). Steps towards an Ecology of Infrastructure: Complex Problems in Design and Access for Large-Scale Collaborative *Systems. In CSCW'94: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*. New York: ACM Press, 253–264.

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

Stengers Isabelle (2018). *Another Science Is Possible: A Manifesto for Slow Science*. Cambridge: Polity.

Strathern Marilyn (2005a). Experiments in interdisciplinarity. *Social Anthropology* 13(1), 75–90.

Strathern Marilyn (2005b). *Kinship, Law and the Unexpected: Relatives Are Always a Surprise*. New York: Cambridge University Press.

Taylor Charles (1971). Interpretation and the sciences of man. *The Review of Metaphysics* 25(1), 3–51.

Thoegersen Jennifer L. (2018). 'Yeah, I guess that's data': Data practices and conceptions among humanities faculty. *portal: Libraries and the Academy* 18(3), 491–504.
Tsoukas Haridimos (2013). Organization as chaosmos. In Robichaud Daniel (ed.), *Organization and Organizing*. London: Routledge, 52–65.
Vetter Jeremy (2016). *Field Life: Science in the American West during the Railroad Era*. Pittsburgh: University of Pittsburgh Press.
White Hayden (1980). The value of narrativity in the representation of reality. *Critical Inquiry* 7(1), 5–27.
Wilkinson M. D. et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018. https://doi.org/10.1038/sdata.2016.18
Williams Graham (2011). *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*. New York: Springer.
Williamson Oliver E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology* 87(3), 548–577.
Wittgenstein Ludwig (1969). *On certainty*. Oxford: Blackwell.