



---

# An Introduction to Paradata

Lisa Andersson, Isto Huvila, and Olle Sköld

---

## Abstract

To address the challenge of data process and practice descriptions, the aim of this volume is twofold. First, we propose the paradata concept as a method to think about and describe data, information, and knowledge processes and practices. Second, by inviting scholars and practitioners from a wide variety of disciplines, we explore how the paradata concept can be useful in and for information and knowledge management in a wide range of settings. The volume brings together scholars and practitioners from a broad range of subject areas, each offering a distinctive perspective on paradata in different contexts, highlighting diverse scenarios in which collection, extraction, and use of such data may prove constructive. The introductory chapter gives a brief history of the paradata term including insights into current research on the topic. Thereafter follows a concise contextualization of the study of paradata in the knowledge management tradition. The chapter is concluded by a guide to the volume's chapters.

---

## 1 Introduction

A pivotal question linked to data, information, and knowledge is what we need to know about data, information, and knowledge and how it has been managed in order to use it? Sometimes the answer is “very little.” There is, for example, no harm in assuming that a temperature value is exact rather than rounded or not knowing the details of how a particular thermometer works when checking tomorrow's weather

---

L. Andersson · I. Huvila (✉) · O. Sköld  
Uppsala University, Uppsala, Sweden  
e-mail: [lisa.andersson@ki.se](mailto:lisa.andersson@ki.se); [isto.huvila@abm.uu.se](mailto:isto.huvila@abm.uu.se); [olle.skold@abm.uu.se](mailto:olle.skold@abm.uu.se)

© The Author(s) 2024  
I. Huvila et al. (eds.), *Perspectives on Paradata*,  
Knowledge Management and Organizational Learning 13,  
[https://doi.org/10.1007/978-3-031-53946-6\\_1](https://doi.org/10.1007/978-3-031-53946-6_1)

forecast. At other times the situation may be the opposite; the precision of a value is crucial to its usefulness for a particular purpose, e.g., the precision of a temperature value for analyzing global warming. This is where paradata, i.e., information on the processes of data creation, curation, and earlier and intended use, can be pivotal for the usefulness of data, information, and knowledge.

Much of the intricacy of paradata lies in its ubiquity. Like Larry Prusak predicted the future of knowledge management, the knowledge of processes tends to be “thoroughly adopted—so much a natural part of how people organize work—that it eventually becomes invisible” when we engage in our daily pursuits (Prusak, 2001, p. 1006). However, as the temperature example shows, while simply accepting a temperature value can be sensible in many daily tasks, neglecting to account for the process can have major and possibly detrimental consequences in situations where precision is needed. To address the challenge of process and practice descriptions, the aim of this volume is twofold. First, we propose the paradata concept as a method to think about and describe process knowledge, particularly in relation to data, information, and knowledge processes and practices. Second, by inviting scholars and practitioners from a wide variety of disciplines, we explore how the paradata concept can be useful in and for information and knowledge management in a wide range of settings. By this exploration of process knowledge, how process information can support the management of information and knowledge, and of how process information and knowledge can be managed in a variety of contexts we hope to contribute to theoretical and practical advancement in the field of information and knowledge management.

Rather than starting with one particular definition of what paradata is, we embark on an exploration on what paradata can be, both in conceptual and practical sense. As a starting point of this journey, we borrow a definition of “data” from the information science scholar Christine L. Borgman stating that data refer to “entities used as evidence of phenomena for the purposes of research or scholarship” (Borgman, 2015, p. 29) and expand it to include other practices of creating knowledge or knowing also beyond scientific and scholarly pursuits. Further we start off with a working definition of the concept of paradata as “data on processes” following the seminal use of the term in survey research (see the chapter “Paradata in Surveys”). The working definition of paradata serves as a common ground throughout the chapters and creates a space for conceptual exploration of what paradata can be in different settings, the character of the processes they are meant to describe, the methods by which paradata are found or generated, what paradata can do or enable, and what needs to be considered when creating and using paradata in different settings and for different purposes. In the concluding discussion, we summarize and synthesize the various applications of the term as proposed by the 11 chapters.

The volume brings together scholars and practitioners from a broad range of disciplines, each offering a distinctive perspective on paradata in different contexts, highlighting diverse scenarios in which collection, extraction, and use of such data may prove constructive. The heterogeneity of the chapters included is by design; by juxtaposing different approaches to process data and engaging with them,

we spur a wider discussion on the need to record—or reconstruct—data creation processes in order to uncover and make visible previously disregarded and invisible aspects of the creation, curation, and use of the many tangible and intangible resources and materials from which we gather data. Thus, while the theoretical and practical exploration of paradata is rooted in the information and knowledge management tradition, the volume’s inquiry into process data and its management implications branches out into the plethora of fields with which the volume engages. To exemplify, the chapter “Paradata as a Tool for Legal Analysis: Utilising Data-on-Data Related Processes” on paradata in the legal settings demonstrates how paradata can be necessary not only to describe how a piece of evidence came into being, but also to evaluate its evidentiary status in the legal process. As another example, the chapter “Paradata in Emergency Services Communications Systems” on research using emergency service communications data makes a note on two types of paradata, as something created in the course of research activities to make the research comprehensible and in parallel as something sought by researchers to understand the datasets they use. At its outset, rather than inquiring into paradata as a matter of concern for information and knowledge management only, the volume points to the multiple and varied uses of paradata in different scholarly and professional practices to provide insights into diversity of perspectives to paradata and to the intersections of the diverse approaches stemming from a multitude of frames of reference.

Through delving into paradata from the plethora of disciplinary perspectives included in this volume, the chapters forward the understanding of and relevance of paradata as a topic of interest for information and knowledge management. Building on the tradition of theorizing and developing metadata to serve resource description purposes, the inquiry into paradata prompts several leaps forward into new theoretical and practical challenges to engage with. One of the key issues in the earlier paradata-related literature, as well as throughout the chapters of this volume, is how paradata can help to make cognizable and draw attention to doings that underpin different types of information and knowledge. In this sense, the concept points directly to one of the thorniest and most prominent knowledge management problems—that of how to deal with tacit and implicit aspects of knowledge and bring them together with explicit and inscribed forms of information and data (Polanyi, 1967). At the same time, paradata also meanders somewhere in between the two major perspectives to knowledge management as either a social and organizational issue of mediating and translating knowledge between people or a technical issue of inscribing and managing knowledge in, often, technical systems (Handzic, 2004). Depending on how paradata is conceptualized, it can feature as a translational device from either of the two perspectives to bridge the gap in between. Moreover, because paradata often serves to enable a particular data reuse purpose, paradata opens up for theorizing and developing practical solutions to span the spectrum between general, definitive resource description and process information created for situated and time-specific data reuse needs. As with the initial temperature example, an existing set of temperature data can at a later point in time be enriched with thermometer configuration data, if and when the data is

about to be aggregated with other temperature datasets for a more comprehensive analysis.

On a practical note, besides informing practice in the disciplines represented in the individual chapters, the volume will be useful for information and knowledge management professionals seeking to develop paradata documentation in various practices. Also professionals working specifically with data management such as data stewards, data curators, data managers, and data librarians will find the various chapters useful both for drilling deeper into process descriptions in particular disciplinary practices and for understanding the breath and differences between various fields of research and practice. Even information architects and system developers designing data repositories and services for data discovery and reuse will find the volume useful for similar purposes. For researchers and master's and doctoral students the volume will offer a comprehensive overview of a fast-growing field of study with significant and broad relevance for data, information, and knowledge management. Moreover, anyone creating or dealing with data in their daily work will find the volume as food for thought when reflecting on how to describe data processing in a meaningful and efficient way.

This introductory chapter will in the following give a brief history of the paradata term including insight into current research on the topic. Thereafter follows a concise contextualization of the study of paradata in the knowledge management tradition. The chapter is concluded by a guide to the volume's chapters.

---

## 2 The Origins and Status of the Term "Paradata"

The "para-" prefix (Definition of PARA, [n.d.](#)) means beside or alongside of, and thus would make paradata mean information alongside of data. Confusingly, the more established term "metadata" can also mean data about data (Pomerantz, [2015](#)). In practice though, metadata often refers to a condensed and structured description of a resource, commonly guided by a standard agreed upon by some form of community. A well-known example of a metadata standard is the Dublin Core standard developed to describe networked electronic information objects, encompassing categories like "Title," "Type," "Subject," etc. (DCMI Metadata Terms, [n.d.](#)). In order to understand what paradata can bring to the table in terms of data descriptions, it is vital to understand what metadata traditionally does not provide data creators and users a chance to discern and do.

Paradata, as applied for example in statistics to describe survey data, implies a description of the survey process (Couper, [2000](#); for further description, see the chapter "Paradata in Surveys," Paradata in survey research). Similarly, in heritage studies, paradata has been used to denote information on data creation and processing (Denard, [2012](#); see also the chapter "A Leap of Faith: Revisiting Paradata in 3D Scholarship," dismantling the black boxes of 3D scholarship). Even if there is obvious overlap between what could count as metadata and paradata both in theory and practice, in the light of the earlier literature, the paradata notion has a unique and developing conceptual space. Paradata points to a need to document

data processing beyond what is traditionally captured in structured metadata. As a phenomenon, paradata differs from metadata qualitatively in that it seeks to cover the creation and processing of data and quantitatively in that it encompasses more detailed information about data than traditional metadata. Also, according to how it has been conceptualized, the engagement with paradata disrupts established data description standards and practices as data processing information to a large extent is unstructured and not codified. As a positive consequence of that paradata is not (yet) formalized to the same extent as metadata, the previous research and the chapters of this volume show how paradata also unfolds as a potentially powerful device with which to think and explore matters tied to how doings and processes can be documented and understood across domains.

Parallel to paradata, there are other terms partly overlapping in meaning. “Provenance” is a concept common in the archival sphere to describe both creation and curation of informational objects, specifically records. It can encompass both the agents involved and the actions they take, and be useful for acquiring, arranging, retrieving, and appraising records (Sköld et al., 2022). Consequently, and as this edited volume illustrates, scholars and practitioners in information science and knowledge management are not alone in grappling with the challenge to share and understand information on why a data source came to be the way it is. Yet, as of today, the character of and need for paradata remains open for exploration in most fields of data creation. Likewise, the question of its relation to metadata and the challenge of incorporating processing information into structured resource descriptions is in the stage of experimentation. There are well-defined models that cover specific types of paradata, like the CIDOC-CRMdig (Doerr et al., n.d.) standard for documenting the steps and methods of producing digital objects, PROV specification for modeling data provenance (PROV-Overview: An Overview of the PROV Family of Documents, 2013), and Common Workflow Language for inscribing computational analytical workflows (Amstutz et al., 2016). While these models meet parts of users’ paradata needs, a recent analysis of the use of the paradata concept in archaeology and heritage studies points to a number of uncertainties pertaining to the paradata concept that require further study to be clarified. These uncertainties include the challenge of determining the required types and amounts of paradata, considerations of paradata users, and what kind of transparency paradata can be expected to facilitate (Sköld et al., 2022).

The next section takes a closer look at the exploration of paradata in the field of knowledge management and how the emerging interest in paradata both taps into traditional trajectories and opens up new paths for knowledge management theory and practice.

## 3 Paradata in Information and Knowledge Management

### 3.1 Approaching Paradata

The exploration of paradata as seen in this volume and in multiple research literatures has apparent affinities with information and knowledge management, a field of research and practice that concerns the creation, capturing, organization, access, and use of sources of information and knowledge. In this volume, we refer to information and knowledge management as the broad field of research and practice that comprises and is termed in the literature varyingly as knowledge management or information and knowledge management with a focus on managing knowledge, information and data, records, and collections. This framing acknowledges the diversity of the field both regarding differences in research traditions and their analytical focus, including the differences in referring to knowledge management, organizational learning, information management, data management, and records management (Pun & Nathai-Balkissoon, 2011; Schlögl, 2005). However, for us in this volume, similarly to how paradata has been discussed in relation to metadata so far, the boundaries between “information” and “knowledge” are definitional rather than strict. According to both the widely adopted and criticized data–information–knowledge hierarchy, information is commonly understood as a message, and knowledge as information that is contextualized in the minds or practices of knowing individuals or organizations (Davenport, 1998). While previously much of the information and knowledge management literature has focused on one of the two or both concepts, interestingly from the perspective of paradata, lately also the terms “data” and “data management” have been related to and included in the wider field of knowledge management (Dalkir, 2023).

Even if much of the knowledge management literature relates to business contexts referring to the management of knowledge for the efficacy and profit of the business organization (Bolisani & Bratianu, 2018), it has been used in a wider sense as denoting the dealing with information or knowledge as a means to an end—whether it be personal or professional, for leisure or profit. Knowledge management is a thriving topic in the context of library management (Shropshire et al., 2020), and there is a growing corpus of work on personal everyday life information and knowledge management (Dinneen & Julien, 2020; Pauleen & Gorman, 2011; Swigon, 2011). An earlier book in the International Association for Knowledge Management series *Knowledge Management and Organizational Learning*—where this volume is appearing—focused on knowledge management in and in relation to arts and humanities (Handzic & Carlucci, 2019). By inquiring into paradata, we are continuing this interdisciplinary push toward expanding the horizons of information and knowledge management research and practice. Knowledge management research generally attempts to theorize beyond the specific case or cases, using empirical evidence to generate models with a more general explanatory reach. Yet, at the same time it recognizes the cultural, social, and cognitive aspects of and influences on managing knowledge and information. Much similarly to how information and knowledge management has many practical and

theoretical applications in arts and humanities, and arts and humanities perspectives can inform the development of information and knowledge management theory and practice, the multidisciplinary engagement with paradata as put forth by this volume has implications in the affected disciplines and on paradata as an information and knowledge management concept.

Just like how information and knowledge management commonly concerns the system level management of information and knowledge also paradata unfolds as a potentially comparably systemic concept. “System” here refers to technical systems for recording, storing, and sharing information, but also the organizational and social system governing information. The knowledge in question can be both that can be codified and stored in a technical system and the lived knowledge that people know by experience and act on by default. Similarly to a part of the previous and ongoing paradata research, along the lines of the technical strand of information and knowledge management, to reach its goal of efficient information and knowledge sharing, much of the research effort has been put into understanding how knowledge can be codified, shared, and retrieved by support from technical systems. Yet, interdisciplinary work on both paradata and information and knowledge management alike also takes interest in the knowledge sharing taking place beyond technical systems—both to understand the flaws of technical systems in order to improve their functionality and to understand what knowledge cannot be codified and transmitted via a technical system. A now classic study of customer support service operators shows the supremacy of sitting next to the most knowledgeable co-worker, rather than using the knowledge management system provided (Orr, 2016). Correspondingly, the human and the lived experience as sources of paradata should not be overlooked, as evinced by multiple contributions to this volume like the chapter “Dustings of Paradata as Pedagogical Support at Four Archaeological Field-School Sites” on paradata derived from analytical narration of a fieldwork experience and the chapter “Towards Embodied Paradata. A Diffractive Art/Archaeology Approach” on what we can learn about processes from the bodily and embodied information.

### 3.2 Data Descriptions and Knowledge Management

Somewhat counterintuitively, even if the work on taxonomies and development of knowledge management systems have acknowledged and addressed to a degree the significance of organizing resources, the information and knowledge management discipline has never put the description and organization of knowledge in its immediate focus. Therefore, it is perhaps unsurprising that the conceptual work on process descriptions and paradata stems from other disciplinary contexts. Nevertheless, how sources are described and organized is pivotal to their findability and usability in any system also from the information and knowledge management perspective.

In contrast to the relative lack of emphasis of descriptions and the work of describing and organizing, the description and organization of sources is the key focus of the neighboring discipline of knowledge organization (KO) (Smiraglia,

2014). Knowledge organization research deals with metadata as a theoretical construct and practical implementation to make knowledge, information, and data findable, accessible, interoperable, and reusable (Wilkinson et al., 2016). The idea of knowledge, information, and data reusability is the intersection where the interests of knowledge organization and information and knowledge management merge. For a knowledge management system to meet its objectives, the knowledge within needs to be not only findable but also reusable, and reusability presupposes some extent of understanding of what is being made findable and reusable both in terms of their whatness and their processual origins.

But resource description has never been easy. Both knowledge organization and knowledge management have since long acknowledged the complexity of *organizational* as opposed to *individual* knowledge, the *cultural* and *social* contexts of knowledge, and the *tacit* versus the *explicit* components of knowledge. However, the challenges of describing and managing resources differ depending on the type of resource in focus of the description and management. The resource might be as different as a piece of “know-how,” the solution to a common problem traded between software developers writing code, or a static fact about the head count of a customer organization. However, when the resource is “data”—be it a continuous data lake or a bounded dataset, expected to be reusable and have sustained usability over time, the resource presents a new set of descriptive challenges connected to how the data came into being, was organized, and has been managed over time. This is where paradata has the potential to support knowledge management.

While traditional metadata aims at describing sources by assigning attributes like “language” or “date issued,” and sometimes even gives insight into the history of a source by stating “provenance” indicating changes in ownership or custody (DCMI Metadata Terms, n.d.), the rationale behind expanding traditional metadata with additional information on the origin and processing of data is that users likely will need to know more than core metadata in order to understand a data resource to the extent that it becomes usable (Börjesson et al., 2022). Thus, to meet the knowledge management goals about efficient retrieval and reuse of knowledge in the form of data, the data resources need to be sufficiently well-described—often beyond what is achieved by traditional metadata.

Thus, to understand paradata as a rising topic of practical importance and information and knowledge management research interest we need to look at the resource in focus of the management efforts. From a narrow perspective, organizations aim to manage, reuse, and capitalize off of their data assets, and from a broader perspective, societies need to manage their data in sustainable ways to prevent waste of public resources and build critical knowledge, e.g., about demographics and public health, over time. Furthermore, from a research perspective—as the datasets become larger and the methods more advanced—researchers need refined ways of describing their methods to maintain methodological transparency and produce reliable results, as is apparent for example in the case of genetic epidemiology in the chapter “Making Research Code Useful Paradata.” The following final section of this introduction gives a thematic overview of the chapters offering insights into their contributions to the emerging field of paradata research.



## 4 Thematic Overview

While this volume collects contributions concerning paradata from a range of fields, it does by no means offer an exhaustive overview of paradata types and paradata use cases. We hope and believe that giving a more comprehensive overview can be a task for further research. The selection of fields and practices covered in this volume is based on us, the editors, identifying contexts and practices where process descriptions are needed and, in some instances, already explored and practiced. Some, like survey research (chapter “Paradata in Surveys”) and heritage visualization (chapter “A Leap of Faith: Revisiting Paradata in 3D Scholarship”), had a given place, as those are fields where the paradata term has been used already for several decades. Others, like archival science (chapters “Mapping Accessions to Repositories Data: A Case Study in Paradata” and “The Role of Paradata in Algorithmic Accountability”) and legal analysis (chapter “Paradata as a Tool for Legal Analysis: Utilising Data-on-Data Related Processes”), were identified as domains where process descriptions are needed to complete core tasks of the fields. We reached out to scholars in these areas asking them to describe and analyze methods for process descriptions within their field or area of expertise and to use the notion of paradata to discuss the process description and documentation. As a result, the volume combines contributions from authors familiar with using the paradata concept already before writing their chapters and contributions from authors thinking with and describing process descriptions as paradata for the very first time. While we provided the starting point for a conceptually driven exploration, the analyses and reflections are wholly attributable to the chapter authors.

As a result of our highly purposive sampling method, you will meet authors from a wide range of fields and practices, bringing their disciplinary styles and stylistic preferences to the table. While the sample of disciplines and authors is by no means systematic or meant to be such, all chapter authors do, to some degree, engage either explicitly or implicitly in information and knowledge management in their writing—although many have never been aware of or related their work to information and knowledge management theory and research previously. Sometimes the links are obvious, sometimes less so. In the latter case, rather than forcing the chapters to take steps outside of their home domains, we have made an effort to reflect on the links in the final chapter of this volume (Huvila, Andersson, & Sköld, this volume). Moreover, the volume covers chapters on processing descriptions in relation to public as well as private data, research, and governance data. The purposes for storing and processing data vary accordingly, from commercial purposes to heritage preservation.

Naturally, the quest of making process and practice information and knowledge more explicit and more tangible speaks to contemporary, sometimes highly ideological, discourses of openness and transparency. These are familiar from the open science movement (Open Innovation, Open Science, Open to the World | Shaping Europe’s Digital Future, 2016), efforts to make the use of personal data

more transparent, such as the EU's GDPR (General Data Protection Regulation (GDPR), [n.d.](#)), and call for insight into the algorithmic processes shaping our lives (see for example the chapter "The Role of Paradata in Algorithmic Accountability"). Paradata can thus be drawn upon to promote positive values such as accountability, trust, and sustainability in information and knowledge management processes. At the same time, the notion of paradata raises the challenging questions of to what extent intellectual and social as well as technical processes can be described and understood. The task of creating understandable process descriptions challenges our human creativity, as well as the modes of communication we have available, and the forms of documentation we use to inscribe and transmit the communication. This way, the notion of paradata is both an expression of a quest to describe and manage process information and a humbling reminder of how difficult it is to capture process knowledge.

As already noted, the chapters of the volume span between paradata created in research, either as a support for the research process as in the case with computational code in the chapter "Making Research Code Useful Paradata" or as more of a by-product of a study, as exemplified by the chapter on surveys, the chapter "Paradata in Surveys." Also, we see a breadth in the reasons for documenting paradata. Process documentation for transparency of processes for current and future audiences is one motivation, apparent for example in the chapter "A Leap of Faith: Revisiting Paradata in 3D Scholarship" on heritage visualizations and the chapter "Paradata for Digitization Processes and Digital Scholarly Editions" on digitalization processes. Paradata as a direct input into the management, in the sense of control or change of processes, is another motivation, standing out for example in the chapter "Paradata in Emergency Services Communications Systems" on emergency service communication and the chapter "Adding Paradata About Records Processes via Information Control Plans" on public-sector processes.

The volume starts by a chapter by the statistician Patrick Oliver Schenk and the archaeologist Simone Reuß, introducing the field of survey studies where the paradata is a widely adapted concept. The chapter probes deeper into the definition of paradata by comparing the term to the related terms auxiliary data, contextual data, and metadata, and gives ample examples of the variety of methods for collecting paradata. The chapter "Making Research Code Useful Paradata," by the computational biologist Richèl J.C. Bilderbeek, delves into a specific type of paradata, namely computer code in computational research for the purpose of reproducibility. Using the case of genetic epidemiology, Bilderbeek proposes how to improve code to make it serve as paradata. The chapter "A Leap of Faith: Revisiting Paradata in 3D Scholarship," by the digital humanities scholar Costas Papadopoulos, brings us into the practice of 3D (re)construction where the formats for paradata are far more composite. Papadopoulos highlights the variable and dialectic processes of fieldwork and points to the perceptual, physiological, and technical factors that need to be accounted for to understand 3D (re)constructions.

The chapter "Dustings of Paradata as Pedagogical Support at Four Archaeological Field-School Sites," by the archival studies scholar Sarah A. Buchanan and the archaeologist Theresa Huntsman, also concerns paradata from fieldwork,

but focusing on the lived experience as paradata. In their framing, paradata is the collaborative analytical narration, led by a data archivist to support the contextual integrity of the data collected, and to surface the pedagogical goals of each project. The artist and art scholar Ian Dawson and the archaeologist and computer scientist Paul Reilly similarly emphasize the involvement of the human subject in the creation of paradata in the chapter “Towards Embodied Paradata. A Diffractive Art/Archaeology Approach.” Dawson and Reilly propose the notion of “embodied paradata” as a way to understand how the worker and their tools, their bodily practices of making or uncovering knowledge, make up paradata.

The chapters “Mapping Accessions to Repositories Data: A Case Study in Paradata,” “Paradata for Digitization Processes and Digital Scholarly Editions,” and “Reconstructing Provenance in Long-Lived Data Systems: The Challenge of Paradata Capture in Memory Institution Collection Databases” put the paradata concept to work for the purpose of understanding the composition of collections, editions, and databases. In the chapter “Mapping Accessions to Repositories Data: A Case Study in Paradata,” the historian of science Kevin Matthew Jones and the archivist Jenny Bunn explore how paradata can shed light on the choices and assumptions made by archivists in accession processes. In the chapter “Paradata for Digitization Processes and Digital Scholarly Editions,” the information science scholar Wout Dillen discusses how digitized collections often lack the information of how the digitization was made, and thereby leaving the user of the digital reproductions without the necessary clues to understand the quality of the reproduction. Paradata on the digitization process would, in this case, directly improve research validity. The chapter “Reconstructing Provenance in Long-Lived Data Systems: The Challenge of Paradata Capture in Memory Institution Collection Databases,” by information science scholars Alexandria Rayburn and Andrea Thomer, explores how databases’ histories and maintenance can be documented by means of visualizations and thereby made accessible to users.

A number of chapters also apply the paradata concept in situations beyond research and collections development and bring insight into various processes highly definitional of people’s daily lives. The multidisciplinary team behind the chapter “Paradata in Emergency Services Communications Systems” explores how paradata in government archives could support the modeling and simulations needed to try out new technologies for emergency service communications. The information science scholars Ciaran B. Trace and James A. Hodges delve into the role of paradata for algorithmic transparency and explainability of algorithms and algorithmic systems in the chapter “The Role of Paradata in Algorithmic Accountability.” The chapter “Adding Paradata About Records Processes via Information Control Plans” explores how automatically assigned paradata could contribute to information control in public administration. The archival scholars Saara Packalén and Pekka Henttonen discuss how information on the processes from which records originate may further the understanding of the records kept. Lastly, the law scholar Lena Enqvist, similarly to Packalén and Henttonen, explores how paradata could increase transparency in the public sector. Enqvist investigates how paradata could make

public authorities' use of technology in decision making more visible and auditable from a legal perspective.

The concluding chapter draws together insights from the discipline-specific chapters to contrast and synthesize the diverse approaches to how paradata is conceptualized and used. Further, we proceed to three topics of discussion emerging from the synthesizing analysis: how paradata is done in practice, the implications of paradata for the theory and practice of information and knowledge management, and the ethics of paradata. The chapter is concluded by brief remarks on future directions of paradata research and practice. By this exploration of paradata we do, together with the chapter authors in this volume, hope to promote the paradata concept as a valuable contribution to the knowledge management toolbox.

---

## References

- Amstutz, P., Crusoe, M. R., Tijanić, N., Chapman, B., Chilton, J., Heuer, M., Kartashov, A., Leehr, D., Ménager, H., Nedeljkovich, M., Scales, M., Soiland-Reyes, S., & Stojanovic, L. (2016). *Common workflow language, v1.0* (5921760 Bytes) [Data set]. figshare. <https://doi.org/10.6084/M9.FIGSHARE.3115156.V2>
- Bolisani, E., & Bratianu, C. (2018). *Emergent knowledge strategies* (Vol. 4). Springer. <https://doi.org/10.1007/978-3-319-60657-6>
- Borgman, C. L. (2015). *Big data, little data, no data: Scholarship in the networked world*. MIT Press.
- 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>
- Couper, M. P. (2000). Usability evaluation of computer-assisted survey instruments. *Social Science Computer Review*, 18(4), 384–396. <https://doi.org/10.1177/089443930001800402>
- Dalkir, K. (2023). *Knowledge management in theory and practice*. The MIT Press. <https://mitpress.mit.edu/9780262048125/knowledge-management-in-theory-and-practice/>
- Davenport, T. H. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business School.
- DCMI Metadata Terms. (n.d.). Retrieved May 26, 2023, from <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- Definition of PARA. (n.d.). Retrieved August 12, 2022, from <https://www.merriam-webster.com/dictionary/para>
- Denard, H. (2012). *Paradata and transparency in virtual heritage*. Taylor & Francis.
- Dinneen, J. D., & Julien, C.-A. (2020). The ubiquitous digital file: A review of file management research. *Journal of the Association for Information Science and Technology*, 71, E1–E32. <https://doi.org/10.1002/asi.24222>
- Doerr, M., Stead, S., & Theodoridou, M. (n.d.). *CRMDig v4.0*. FORTH. Retrieved May 26, 2023, from <https://www.cidoc-crm.org/crmdig/sites/default/files/CRMDigv4.0.pdf>
- General Data Protection Regulation (GDPR). (n.d.). Retrieved July 5, 2023, from <https://gdprinfo.eu/>
- Handzic, M. (2004). *Knowledge management: Through the technology glass*. World Scientific.
- Handzic, M., & Carlucci, D. (Eds.). (2019). *Knowledge management, arts and humanities*. Springer. <https://doi.org/10.1007/978-3-030-10922-6>
- Open innovation, open science, open to the world | Shaping Europe's digital future. (2016, June 16). <https://digital-strategy.ec.europa.eu/en/library/open-innovation-open-science-open-world>
- Orr, J. E. (2016). *Talking about machines: An ethnography of a modern job*. Cornell University Press. <http://ebookcentral.proquest.com/lib/uu/detail.action?docID=4742043>

- Pauleen, D., & Gorman, G. (Eds.). (2011). *Personal knowledge management: Individual, organizational and social perspectives*. Gower.
- Polanyi, M. (1967). *The tacit dimension*. Routledge.
- Pomerantz, J. (2015). *Metadata*. MIT Press. <http://ebookcentral.proquest.com/lib/uu/detail.action?docID=4397948>
- PROV-Overview: An Overview of the PROV Family of Documents. (2013). <http://www.w3.org/TR/2013/NOTE-prov-overview-20130430/>
- Prusak, L. (2001). Where did knowledge management come from? *IBM Systems Journal*, 40(4), 1002–1007. <https://doi.org/10.1147/sj.404.01002>
- Pun, K. F., & Nathai-Balkissoon, M. (2011). Integrating knowledge management into organisational learning: A review of concepts and models. *Learning Organization*, 18(3), 203–223.
- Schlögl, C. (2005). Information and knowledge management: Dimensions and approaches. *Information Research*, 10(4).
- Shropshire, S., Semenza, J. L., & Koury, R. (2020). Knowledge management in practice in academic libraries. *IFLA Journal*, 46(1), 25–33. <https://doi.org/10.1177/0340035219878865>
- Sköld, O., Börjesson, L., & Huvila, I. (2022). Interrogating paradata. In: *Information research. Proceedings of the 11th international conference on conceptions of library and information science, Oslo Metropolitan University, May 29–June 1, 2022* (Vol. 27, special issue, paper colis2206). <https://doi.org/10.47989/ircolis2206>
- Smiraglia, R. P. (2014). *The elements of knowledge organization* (1st ed.). Springer. <https://doi.org/10.1007/978-3-319-09357-4>
- Swigon, M. (2011). Information limits: Definition, typology and types. *Aslib Proceedings*, 63(4), 364–379.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. -W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

**Lisa Andersson** Ph.D., works as a researcher at the Department of ALM at Uppsala University in Sweden. She received her MA degree in library and information science in 2011, and her doctoral degree in library and information science in 2017, both at the Department of ALM at Uppsala University in Sweden. Her research focuses on data and information management including research data and information management systems, knowledge organization and data descriptions, data publishing, and use. Andersson has published in library and information science journals but also in cross-disciplinary journals from the fields of archaeology and digital humanities.

**Isto Huvila** Ph.D., is professor in information studies at the Department of ALM at Uppsala University in Sweden. He received an MA in cultural history at the University of Turku in 2002 and a Ph.D. degree in information studies at Åbo Akademi University (Turku, Finland) in 2006. Huvila was chairing the recently closed COST Action ARKWORK and is directing the ERC-funded research project CAPTURE. His primary areas of research include information and knowledge management, information work, knowledge organization, documentation, and social and participatory information practices.

**Olle Sköld** Ph.D., works as Senior Lecturer at the Department of ALM, Uppsala University, Sweden. He received his MA degree in archival studies in 2010 at Lund University and a doctoral degree in information studies in 2018 at Uppsala University. His research is characterized by a broad interest in the ALM field, knowledge organization and production, research data creation and use, and digital humanities. Sköld has published in information studies journals including the *Journal of the Association for Information Science and Technology*, *Journal of Documentation*, and *Information Research*.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

