



Archival Practices and the Making of “Memories”

Natasha S. Mauthner & Judit Gárdos

To cite this article: Natasha S. Mauthner & Judit Gárdos (2015) Archival Practices and the Making of “Memories”, *New Review of Information Networking*, 20:1-2, 155-169, DOI: [10.1080/13614576.2015.1114825](https://doi.org/10.1080/13614576.2015.1114825)

To link to this article: <https://doi.org/10.1080/13614576.2015.1114825>



Published online: 17 Dec 2015.



Submit your article to this journal [↗](#)



Article views: 261



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

Archival Practices and the Making of “Memories”

NATASHA S. MAUTHNER

University of Aberdeen Business School, Old Aberdeen, UK

JUDIT GÁRDOS

*Hungarian Academy of Sciences, Center for Social Sciences, Institute of Sociology,
Budapest, Hungary*

Following Derrida (1995), our article explores the relationship between archival practices and archival documents on the assumption that “archivization produces as much as it records the event” (Derrida 1995, 17). On this approach, archival practices are understood as non-innocent practices that, in the act of “preservation,” help make specific “memories” at the expense of others (Barad 2007; Derrida 1995; Foucault 1972). We take up this issue in relation to the curation of social science quantitative research data and argue that the ontological identity of data is constituted through historically- and culturally-specific data curation practices including data cleaning, data anonymization, and metadata preparation.

KEYWORDS *philosophy of the archive, archival science, social science data archives, data curation, data cleaning, data anonymization, metadata preparation*

INTRODUCTION

In both archival and social sciences data curation practices have tended to be understood in technical and administrative terms: as neutral means of safeguarding documents and data, where the objectivity of these practices is secured by following agreed professional standards and protocols.

© Natasha S. Mauthner and Judit Gárdos

Address correspondence to Natasha S. Mauthner, University of Aberdeen, Business School, Dunbar Street, Old Aberdeen AB24 3QY, UK. E-mail: n.mauthner@abdn.ac.uk

Postmodern thinking coupled with technological change have challenged this normative understanding of the archive. Works by Derrida (1995) and Foucault (1972) have been particularly influential, causing archivists and scholars alike to reconfigure archival science and practice to explore: the nature of archives as socially constructed institutions, the role of archives in the production of knowledge about the past, and the power of archives and records to shape our notions of history, identity, and memory (Schwartz and Cook 2002, Cook and Schwartz 2002).

In this article we take up these ideas specifically in relation to the curation of social science quantitative research data. Our argument is that data curation practices do not innocently preserve data but help constitute their ontological identity. Data curation practices are “performative” (Butler 1990; Barad 2007; Law 2004) in that they help bring into being the data they ostensibly preserve. They do so, we suggest, not simply through the subjectivity of data curators, or the technologies and infrastructures of the data repositories; but rather through the specific conceptual assumptions that are embodied and enacted in data curation practices. Data curation and archival practices, we suggest, can be understood as historically- and culturally-specific and contingent “metaphysical practices that necessarily enact specific metaphysical commitments to the exclusion of others” (Mauthner 2015, 331). On this approach, data curation practices are an ineliminable and constitutive part of the data they help bring into being.

We conceptualize data curation practices in broad terms as including specific practices—for example, cleaning up of datasets; data anonymization; streamlining interviews; data storing, categorizing and visualization; and data search tools—as well as a wider range of knowledge, ethical, moral, legal, political, and economic practices these practices are entangled with—for example, field- and discipline-specific knowledge-making practices; national/international data management and curation policies, practices and guidelines; research ethics and governance guidelines and frameworks provided by professional bodies; country-specific data protection, copyright, and information sharing legislation. For the purposes of this article we focus on three specific data curation practices—data cleaning, data anonymization, and metadata preparation—and investigate the ontological processes through which these practices help constitute the survey data they ostensibly archive.

Our article is organized in five parts. We begin with a discussion of Derrida’s *Archive Fever* (1995), and go on to consider how his ideas and postmodern scholarship, more generally, have contributed to discussions about the philosophy of the archive in archival science. We then turn to consider the nature and extent of philosophical debate in relation to social science data archives. In the next section we outline some theoretical ideas around the performativity of knowledge-making practices that inform our own attempts to conceptualize data curation practices in

performative/nonrepresentational terms. In the final part we provide some examples of how data curation practices help *make* the ontological identity of data by arguing that these practices embody and enact specific historically- and culturally-contingent concepts and categories.

JACQUES DERRIDA’S *ARCHIVE FEVER*

In *Archive Fever*, Jacques Derrida (1995) takes the archive as his object of study. He challenges normative understandings of the archive as a neutral and innocent keeper of historical records and artifacts that tell of a past that really happened, and that can be returned to for the purposes of constructing ordinary tales. Derrida troubles this commonsensical view by reminding readers of the etymology of the word “archive.” The term, he suggests, refers to the *arkhē*, which is the source, beginning, origin, in two senses: ontological and nomological. *Arkhē* denotes at once the commencement and the commandment. The meaning of “archive,” Derrida goes on, comes from the Latin *archivum* or *archium*, itself derived from the Greek *arkheion*. The latter was the residence of the *archons*: the superior magistrates who commanded and ruled in ancient Greece. On account of their publicly recognized authority, their home became the repository of official documents. Not only did the archons become the guardians of these documents, but they were also accorded hermeneutic rights: they had “the power to interpret the archives” (Derrida 1995, 3).

Derrida points further to the politics of the very constitution—and not only interpretation—of archives and archival content. Archives, Derrida argues, come into existence in highly specific and inseparably ontological and political configurations, what he calls “a privileged *topology*” (Derrida 1995, 3). It is in this sense that archives both *embody* and *enact* a specific politics: they are both expressions and instruments of power (Harris 2002). In Derrida’s terms, archival “documents in effect, speak the law: they recall the law and call on or impose the law” (Derrida 1995, 3). The processes and practices through which archives are constituted both “lay down the law and *give the order*” (Derrida 1995, 7).

This order—in the form of classificatory concepts and systems—is mutually constitutive of archives and archival records. As Derrida puts it: “the technical structure of the *archiving* archive also determines the structure of the *archivable* content even in its very coming into existence and in its relationship to the future. The archivization produces as much as it records the event” (Derrida 1995, 17). The systems through which archival content is organized/ordered—for example, whether a record is classified as “theory” or “private correspondence,” “biography” or “autobiography”—contribute toward the *making* of archival content. Derrida calls this coming into existence in a specific place and according to a specific law (classificatory

system) the “archontic principle.” The archontic principle of the archive, he suggests, “is also a principle of consignment, that is, of gathering together” (Derrida 1995, 3) where “*Consignment* aims to coordinate a single corpus, in a system or a synchrony in which all the elements articulate a unity of an ideal configuration” (Derrida 1995, 3).

Derrida’s point is not that classificatory concepts and systems can or should be avoided as without them “no archive would ever come into play or exist as such” (Derrida 1995, 3). In order to exist, every archive must necessarily come into being in a specific form. Rather, the issue is that the politics and “violence” (Derrida 1995, 7) through which the archive must be, and is inevitably, constituted is naturalized and taken as given. Derrida wants to make apparent that the archive is a place “where law and singularity intersect in *privilege*” (Derrida 1995). He argues that:

A science of the archive must include the theory of this institutionalization, that is to say, the theory both of the law which begins by inscribing itself there and of the right which authorizes it. This right imposes or supposes a bundle of limits which have a history, a deconstructable history. (3)

Derrida (1995) points to the ways in which archives embody and enact power in two ways: first, through privileged topologies (concepts and categories) that help make the law (specific ontological and political realities); and second, by naturalizing, forgetting, and erasing the fact that these *are* privileged topologies: that archives are a materialization of specific ontological and political commitments to the exclusion of all others. It is in this sense, Derrida argues, that the concept of the archive carries but also forgets the name *arkhē*.

PHILOSOPHY OF THE ARCHIVE

Derrida’s writings (1995) on the archive, along with those of Foucault (1972) and other postmodern thinkers, have fueled debate about “the philosophy of the archive” (Whatley and Brown 2009). For over two decades, archivists and scholars have been grappling with “the new realities being fashioned by technological revolution and postmodernist epistemologies” (Harris 2002, 83), and their implications for the conceptualization and practice of archival work. Both postmodern thinking and the shift to electronic records have challenged normative and naturalized assumptions about the neutrality, objectivity, and given-ness of archives, archival records, and archival practices. The shift to electronic record-making and keeping has helped make visible the hitherto invisible: the role of the archivist and archival practices in the very act of record-creation and preservation, thereby dispelling the notion of the

archivist as innocent guardian of the archival record (Hedstrom 2002; Trace 2002). Similarly, postmodern scholarship has troubled normative technical understandings of archival science, rooted in nineteenth century positivist and administrative-judicial frameworks. As Cook (2013, 100) notes, early literature on archiving (Jenkinson 1947) stresses that the goal was to “preserve the records as evidence of the functional-structural context and actions that caused their creation.” Cook (2013, 106) calls this approach pre-modern, the archivist being a passive keeper of an entire documentary (Cook 2013, 102). Duff and Harris (2002) see the roots of this positivistic archival practice in the traditions of Enlightenment. According to them, the understanding behind these practices is this:

The boundary between text and context is hard and stable. A record’s context is bounded and readily knowable. The archivist’s role in relation to records is to reveal their meaning and significance—not to participate in the construction of meanings—through the exercise of intellectual control. (Duff and Harris 2002, 264)

Under the influence of the postmodern “archival paradigm shift” (Cook 2001), archivists and scholars have challenged “unquestioned assumptions underlying the nature and meaning of archives in society” (Schwartz and Cook 2002, 5). They have destabilized representational approaches in which the archival record is seen to reflect or represent reality and “stand for the facts it is about” (Duranti 2001; see also MacNeil 2004). There has been a shift away from treating the archive-as-source to taking archives, archival records, and archivists as objects of study in their own right (Stoler 2002).

The almost exclusive focus of archival science on the technology and mechanics of archival processes has given way to consideration of “what archives, records, and archivists do on a philosophical or theoretical level, the power they wield, the impact they have” (Schwartz and Cook 2002, 18–19; Cook and Schwartz 2002). A major focus of investigation has been on power (Cook 2001). Archivists and scholars have investigated how power is exercised both through the specificity of the cultural frameworks that shape archival records, archives as institutions and archival practices, and through the *naturalization* of these implicit underpinning cultural frameworks. Archival records and archives are no longer viewed as naturally (given) constructs but as socially constructed. The research focus has shifted to the social, historical, cultural, and political power constellations, as well as to the standards, values and ideologies that shape the creation and maintenance of archives, archival records, and the archival profession (Trace 2002).

This postmodern archival practice, argues Cook, takes the form of social activism for “memory-meaning, adopting a flexible, fluid, and pluralistic mentality mirroring the values of postmodern society and the possibilities of

digital technology” (Cook 2013, 111). In parallel to this, however, and bound to a modernist tradition of archiving, archivists are also constantly developing more sophisticated means by which archives are managed and evidence is protected, with a focus on creating and implementing standards, record-keeping requirements, process templates, and system architectures. Here, rigid consistency of professional practice is sought. Such work reveals, Cook (2013) suggests, the continuing concern for evidence among the memory-dominated and identity-formation paradigm. “As a result, between the poles of evidence and memory, there was sometimes considerable tension in professional discourses, between ever more sophisticated and complex modernist techniques for evidence protection reflecting a culminating expertise in that regard and ever more contextualized and contingent postmodern ideals in turn reflecting contemporary societal values” (Cook 2013, 111). He sees the many contemporary archivists’ endeavors around electronic data preservation in the same old modernist line of tradition (Cook 2013, 100–01) to preserve the data and how they came into being. As we go on to discuss, in the context of social scientific data archiving, such modernist practices can involve state-of-the-art metadata schemes and archival software products.

PHILOSOPHY OF SOCIAL SCIENCE DATA ARCHIVES

As in the field of archival science, philosophical issues in relation to social science data archives have been discussed since the late 1990s. Theoretical and methodological engagement and uptake of these ideas, however, has been limited with little attention given to the philosophy and politics of social science data archives (Mauthner and Parry 2009, 2013; Mauthner 2015; Mottier 2005). Few systematic critical and/or ethnographic investigations have been undertaken exploring how the processes and practices through which social science projects and data are archived shape the resulting archived research collection. In contrast, there are many texts outlining normative data archiving and curating principles and practices and their underpinning ethical, legal, governance, scientific, and technical frameworks and infrastructures. Poole (2014), for example, discusses the infrastructure of science data curation, including the roles of cyberinfrastructure, research communities, collaboration, planning, policy, and standards and best practices. He also addresses the role of institutions—archives, research libraries, institutional repositories, and centers—in curating science data and the role of archival principles, such as provenance, selection and appraisal, authenticity, metadata, risk management, and trust in digital curation.

Overall, discussions regarding the archiving of social scientific data are rooted within “traditional archival thinking” and a “solely technical framework” (Trace 2002, 150) that overlooks the philosophical dimensions of processes of record creation and preservation. As Mauthner and Parry (2013,

56) suggest, “The focus has . . . been on improving methodological, ethical, legal, scientific and technological infrastructures. For example, ethical issues have been addressed through the development of more sophisticated ethical protocols for data reuse which are understood to ensure better compliance with ethical guidelines (regarding informed consent and data anonymization) and legal requirements for data sharing, and restriction of data reuse to ‘bona fide’ researchers (. . .). Similarly, the formulation of data standards and the requirement to archive contextual information (e.g. through metadata, hyper-text or hypermedia) have been seen as important in rendering “raw” data more meaningful by increasing their representational accuracy (. . .).”

The assumption within these approaches—that data can be treated independently of their ontological contexts of production—has been seen as particularly problematic for qualitative researchers influenced by the postmodern turn and working within interpretive and social constructivist traditions for whom data are understood as reflexively constituted through historically- and culturally-specific practices (Hammersley 1997, 2010; Mauthner et al. 1998; Parry and Mauthner 2004; Savage 2005). These philosophical concerns have been tackled through practices seeking to ensure that contextual information is archived alongside the data. This has been understood as rendering data (and resultant knowledge) more meaningful by enabling researchers to better understand the conditions through which data are generated (Corti 2011). The effect of these practices is to make possible the creation of a new object of study: data and the contexts that constitute them, rather than data alone (e.g., Savage 2005, 2010; Moore 2006; Thomson 2014). However, as Mauthner (2015) argues, the shift to taking context into account still conceptualizes data archiving in the same technical terms in which data curating is understood as a neutral means of preserving data *and* context. These practices therefore continue to enact an “ontology of given realities” (Mauthner 2015, 234)—such as data and context—by overlooking the *constitutive* nature of data archiving and curation practices. Data and context are still taken to be ontologically separate and separable, rather than mutually constitutive.

A PERFORMATIVE CONCEPTUALIZATION OF KNOWLEDGE-MAKING PRACTICES

This enactment of data archiving and curation practices relies on largely hidden and commonsense *representational* assumptions about the nature of reality in which the world is understood as given, and the practices of knowing (e.g., preserving, archiving, curating) are bracketed out and treated as technique (Law 2004). Barad (2007, 53) explains that representationalism “marks a failure to take account of the practices through

which representations are produced.” A representational conceptualization of knowledge-making practices (e.g., data curation practices), she suggests, “takes the notion of separation as foundational. It separates the world into the ontologically disjunct domains of words and things, leaving itself with the dilemma of their linkage such that knowledge is possible” (Barad, 2007, 137). Drawing on Rouse (1996), Barad further argues that representationalism underpins both empirical realism and postmodern philosophical approaches that turn to language and discourse, as both share the representational belief that knowledge mediates access to the material world (reality). Where they differ is on what they take to be their referent: whereas realist claims are understood to represent things in the world as they really are (i.e., nature), social constructivist ones are seen to represent objects that are the product of social activities (i.e., culture). Moreover, attempts to acknowledge the knower through reflexive approaches are also founded on representationalism in that they take for granted the notion that representations reflect (social or natural) reality. Reflexivity, Barad suggests, still holds the world at a distance: it “is based on the belief that practices of representing have no effect on the objects of investigation and that we have a kind of access to representations that we don’t have to the objects themselves” (Barad 2007, 87).

On our account, current configurations of data curation practices, and the wider assemblage of practices they are part of, are being enacted on implicit representational terms because their constitutive role in the formation of data is overlooked. Our approach in this article is to conceptualize data curation practices along nonrepresentational lines, through a philosophical framework that is able to materialize the constitutive effects of data curation (and wider) practices on their objects of study and knowledges produced. For this, we turn to Karen Barad’s (2007) posthumanist performative metaphysics, a framework that embodies and enacts a non-classical ontology in which entities are not taken as given but as constituted through material-discursive practices. Barad’s work challenges classical—Newtonian and Cartesian—metaphysics and provides a new ontology, epistemology, and ethics that she terms “agential realism.” On her account, knowledge-making practices are an ineliminable and constitutive part of the realities they help bring into being. On our reading, her scholarship provides a distinctive metaphysical framework that can materialize, and help reconfigure, the representational ontological assumptions that are embedded and enacted in data curation practices and the wider regime of practices of which they are part.

Barad’s framework (2007) is being taken up in the social sciences, with scholars developing different theoretical and methodological applications of agential realism including ways of approaching data archiving and revisiting (e.g., Tamboukou 2014). In this article we expand on Mauthner’s approach to Barad’s framework, and her broader program of work that seeks to develop a nonrepresentational—posthumanist performative—conceptualization and enactment of research methods and knowledge-making practices in the

social sciences. In particular, Mauthner argues that knowledge-making practices (such as data curation) make realities through the specific meta-physical and conceptual assumptions these practices embody and enact (see Mauthner 2015; Mauthner “The Listening Guide;” Mauthner, “Un/re-making Method”).

THE PERFORMATIVITY OF SOCIAL SCIENCE DATA CURATION PRACTICES: SOME ILLUSTRATIVE EXAMPLES

In this section we explore data curation practices and the conceptual assumptions underpinning these practices. We draw on Judit Gárdos’ work at the Research Documentation Center (RDC) of the Hungarian Academy of Sciences’ Center for Social Sciences where she works as part of a team of social scientists curating data from social science research projects. Their curation practices are informed by the Data Documentation Initiative-Lifecycle (DDI-L, <http://www.ddialliance.org/Specification/DDI-Lifecycle/>).¹ For the purposes of this article we focus our attention on the curation of social science survey data. We limit ourselves to three specific data curation practices: data cleaning, data anonymization, and metadata preparation. Following on from our discussion of performative approaches to knowledge-making practices, our interest is in exploring the conceptual assumptions underpinning these data curation practices.² Herein, we build on insights from the field of archival science and the work of archivists and scholars, such as Schwartz and Cook (2002), who discuss the alleged value-free tools of archival practice—such as standards and templates—and how these impose their own rational, systematic way of seeing on a world of records, record keeping, and records creators (see also Trace 2002; Hedstrom 2002).

Data Cleaning

Data cleaning involves tidying up survey values that appear not to make sense, are mislabeled, or are not provided. These practices are designed to address the problem of “missing data,” a term derived from statistical research which refers to instances where no data value is stored for the variable in an observation either because no data value was recorded or because the reader is unable to make sense of the data value provided. Some data values, such as 1964 or 1981 for “year of birth,” are self-explanatory. In other cases, however, the data values require explanation, such as “1” and “2” in a

¹ The DDI-L is similar to the Digital Curation Center’s Curation Lifecycle Model (see Higgins 2008).

² Data are made by wide-ranging practices, including data collection and data analysis practices, and their underpinning conceptual assumptions. The focus of our article, however, is specifically on data curation.

variable labeled “sex.” In the case of gender where the value “1” is labeled male and “2” is female we may find values of 3, 4, 9, or 99. In such cases, we usually convert these into missing values. The original values remain visible but will be excluded from statistical procedures. Following the DDI-L standard, and its aim to document every step of the data curation process, we add in the curation documentation how we proceeded with such values.

While such data cleaning practices are normatively viewed as neutral and technical, on a performative approach they are understood as helping to constitute the ontological identity of data: both what data are and what they mean. The purpose of data cleaning practices is to make every value of every variable understandable for researchers: to create a dataset that is meaningful, intelligible and usable by imagined and anticipated social scientists and social science research. Yet, although there are technical guidelines for dealing with missing data—the UK Data Archive (UKDA) for example informs data curators to look for “unlikely or impossible values for interval variables”—there is no overarching paradigm in relation to what might be regarded as “unlikely” or “impossible” (Sana and Weinreb 2008, 515). In practice, as Sana and Weinreb (2008) explains, notions of “likely” or “possible” are enacted in normative, culturally- and historically-specific, terms. Sana and Weinreb (2008, 517) give the following examples: in most settings, a teacher is likely to have more than a primary school education; or “someone who reports no sexual contact in the last year is not likely to be young, currently married, and co-resident with their spouse.” It is therefore easy to imagine cases where data are not faulty but just reporting situations that are normatively understood as “unlikely” or “impossible.” It is in this sense that, as we suggest, data cleaning practices embody and enact normative conceptual assumptions—in this case, about teacher training and sexual conduct—that help to constitute the very nature of the data.

Labeling practices are also informed by social science research practices more generally. For example, quantitative surveys and analysis of datasets typically classify income into 4 to 10 groups, which provides both anonymization of the exact income level and enables standard ways to analyze data. Data curators normally have knowledge of how these income groups are constructed by the researchers. In cases where labeling is absent in a dataset, curators use this prior knowledge to label the variables. Their labeling practices therefore instantiate conceptual categories that are normative and accepted within quantitative social scientific work.

Data Anonymization

Anonymization of research collections that contain personal data (i.e., data that can be connected with a single identifiable person) is a very common archival practice. There are various methods for anonymizing data. Following the UKDA guidelines, the process of anonymization depends on whether the

data are quantitative or qualitative. According to the UKDA, the purpose of anonymizing quantitative data is to protect the identity of respondents while retaining “as much meaningful information as possible.” Their guidelines suggest:

Aggregate or reduce the precision of a variable such as the respondent’s age and place of residence. As a general rule, report the lowest level of geo-referencing that will not potentially breach respondent confidentiality. The exact scale depends on the type of data collected, but very detailed geo-references like full postcodes, wards or names of small towns or villages are likely to be problematic. (UK Data Archive)

In designating particular pieces of information for deletion or retention, data anonymization practices enact specific but implicit conceptualizations of both “the personal” and “the meaningful.” These conceptualizations are historically- and culturally-contingent notions that find expression in data curation guidelines and practices, as well as in the field of social sciences more generally. In this sense, data anonymization is not a neutral practice deleting “personal” information or retaining “meaningful” information as if these categories are given and self-evident. Rather, data anonymization helps to constitute the ontological nature and identity of “the personal” and “the meaningful.”

Metadata Preparation

Metadata are defined and understood as data that describe and give information about other data. They are seen to provide contextual information to facilitate “meaningful” interpretation about the data and are viewed as being as important as the data themselves. To illustrate metadata we have randomly chosen a dataset on “Election Funding of Finnish MPs 2007” accessed from the Consortium of European Social Science Data Archives (CESSDA) data catalogue [see Appendix 1]. On the left of the image a list of “variable descriptions” is provided under “Metadata.” One of these, highlighted on the right of the image, is “The amount of contributions received from party organizations (in euros).” Some basic statistics are also provided such as the number of “valid cases.” We make two observations about metadata preparation practices, their underlying assumptions, and their performative effects.

First, the categories and concepts of metadata—the variables listed such as candidate’s name, gender, year of birth, political party, and so forth—are not neutral descriptors but rather help constitute data in specific ways according to classificatory systems that are currently normative, accepted and regarded as “meaningful.” Indeed, metadata schemes have been standardized (e.g., according to DDI-L scheme) because lack of standard formats is seen

to hamper the ability to search across studies for similar items or similar studies (Rasmussen and Blank 2007, 62). Through their classificatory systems, metadata practices act to institutionalize and further solidify these systems, systems that are specific ways not only of understanding but of *making* the world. Echoing Derrida's (1995) aforementioned points, the power of these classificatory systems derives from their claim to innocence; from the notion that metadata are neutral descriptors of data.

Second, increasingly researchers use metadata as a portal into, and way of navigating, a dataset. Metadata enable researchers to interpret variables without having to access the whole dataset. Such a practice is not encouraged by CESSDA, as indicated at the bottom of the table in Appendix 1: "The frequencies displayed are unweighted. All results need careful interpretation. Original data collectors, depositors and the data archive bear no responsibility for any results or interpretations arising from the secondary use of the data." Nevertheless, when researchers use metadata as an entry point into data they are working with data that have been (re)made through metadata preparation practices. Indeed, they are working with data that is being constituted through multiple practices: including but not restricted to data generation, data curation, and data analysis. Although these practices are taken for granted and normatively treated as neutral, we argue that they all help constitute data and metadata. Neither data, nor metadata, nor the distinction between them, are ready-made entities. Rather, data, metadata, and the boundary between them are made through historically- and culturally-specific and contingent data curation practices including data cleaning, data anonymization and metadata preparation.

CONCLUSIONS

Normative practices in archival and social science treat archived data, records and artifacts as representations of reality. We have sought to challenge this view through an examination of specific data curation practices in the social sciences. Archived data, we suggest, are made and remade through multiple practices including data generation, data curation and data analysis. In particular, we have suggested that data curation, and other, practices make data through the conceptual commitments they embody and enact. Data curation (and other) practices are normatively treated as neutral and innocent techniques. This conceptualization of data curation practices renders invisible their constitutive role in the ontological formation of data, records and artifacts. Following Derrida, however, the power of archival and data curation practices lies in their bringing together, and materialization of, a conceptual configuration that is "at once visible and invisible" (Derrida 1995, 3): that is both specific and naturalized. As others have argued (e.g., Somers 1996,

2008) social science has a role to play in opening up these conceptual configurations for empirical and theoretical debate and investigation. In failing to do so, Harris (2002, 84) suggests, "we deny our audience the very space in which democracy thrives."

REFERENCES

- Barad, K. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham, NC: Duke University Press, 2007. Print.
- Butler, J. *Gender Trouble: Feminism and the Subversion of Identity*. London: Routledge, 1990.
- Checkland P., and S. Holwell. *Information, Systems and Information Systems: Making Sense of the Field*. Chichester: John Wiley & Sons, 1998. Print.
- Cook, Terry. "Archival Science and Postmodernism: New Formulations for Old Concepts." *Archival Science* 1 (2001): 3–24. Print.
- . "Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms." *Archival Science*, 13 (2013): 95–120. Print.
- Cook, T., and J. M. Schwartz. "Archives, Records, and Power: From (Postmodern) Theory to (Archival) Performance." *Archival Science* 2.3 (2002): 171–85. Print.
- Corti, L. 2011. "The European Landscape of Qualitative Social Research Archives." *Forum: Qualitative Social Research* 12, Art. 11 (2011). Web. 21 May 2015. <<http://www.qualitative-research.net/index.php/fqs/article/view/1746/3248>>.
- Derrida, J. *Archive Fever*. Chicago: University of Chicago Press, 1995. Print.
- Duff, Wendy M., and V. B. Harris. "Stories and Names: Archival Description as Narrating Records and Constructing Meanings." *Archival Science* 2 (2002): 263–85. Print.
- Duranti, L. "The Impact of Digital Technology on Archival Science." *Archival Science* 1 (2001): 1, 39–55. Print.
- Foucault, M. *The Archaeology of Knowledge and the Discourse on Language*. New York: Pantheon Books, 1972. Print.
- Hammersley, M. "Qualitative Data Archiving: Some Reflections on Its Prospects and Problems." *Sociology* 31 (1997): 131–42. Print.
- . "Can We Re-use Qualitative Data via Secondary Analysis?" *Sociological Research Online* 15. 5 (2010). Web. 15 May 2015. <<http://www.socresonline.org.uk/15/1/5.html>>.
- Harris V. "The Archival Sliver: Power, Memory, and Archives in South Africa." *Archival Science* 2.1-2 (2002): 63–86. Print.
- Hedstrom, M. "Archives, Memory, and Interfaces with the Past." *Archival Science* 2.1-2 (2002): 21–43. Print.
- Higgins, S. "The DCC Curation Lifecycle Model." *The International Journal of Digital Curation* 1.3 (2008): 134–40. Print.
- Jenkinson, H. "The English Archivist: A New Profession." 1947. *Selected Writings of Sir Hilary Jenkinson*. Ed. R. H. Ellis and P. Walne; Allan Sutton: Gloucester, 1980. Print. Issued with new introduction in 2003 by Eastwood TM. Society of American Archivists, Chicago. Print.

- Law, J. *After Method: Mess in Social Science Research*. Abingdon: Routledge, 2004. Print.
- MacNeil, H. "Contemporary Archival Diplomatics as Method of Inquiry: Lessons Learned From Two Research Projects." *Archival Science* 4 (2004): 199–232.
- Mauthner, Natasha. "'The Past was Never Simply There to Begin with and the Future is Not Simply What Will Unfold': A Posthumanist Performative Approach to Qualitative Longitudinal Research." *International Journal of Social Research Methodology* 18.3 (2015): 321–36. doi:[10.1080/13645579.2015.1022298](https://doi.org/10.1080/13645579.2015.1022298).
- . "The Listening Guide Method of Narrative Analysis: Towards a Posthumanist Performative (Re)configuration." *Feminist Narrative Research: Opportunities and Challenge*. Ed. J. Woodiwiss, K. Smith, and K. Lockwood; Basingstoke: Palgrave Macmillan, Forthcoming. Print.
- . "Un/re-making Method: Knowing/Enacting Posthumanist Performative Social Research Methods through 'Diffractive Genealogies' and 'Metaphysical Practices.'" *Mattering: Feminism, Science and Materialism*. Ed. V. Pitts-Taylor; New York: New York University Press, Forthcoming. Print.
- Mauthner, Natasha S., and O. Parry. "Open Access Digital Data Sharing: Principles, Policies and Practices." *Social Epistemology* 27 (2013): 47–67. Print.
- . "Qualitative Data Preservation and Sharing in the Social Sciences: On Whose Philosophical Terms?" *Australian Journal of Social Issues* 44 (2009): 289–305. Print.
- Mauthner, Natasha, O. Parry, and K. Backett-Milburn. "The Data Are Out There, or Are They? Implications for Archiving and Revisiting Qualitative Data." *Sociology* 32 (1998): 733–745. Print.
- Moore, N. "The Contexts of Context: Broadening Perspectives in the (Re)use of Qualitative Data." *Methodological Innovations Online* 1 (2006): 21–32. Print.
- Mottier, V. "The Interpretive Turn: History, Memory, and Storage in Qualitative Research" [21 paragraphs]. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* [On-line Journal], 6.2, 33 (2005). Web. 28 May 2015. <<http://www.qualitative-research.net/index.php/fqs/article/view/456>>.
- Parry, O., and N. Mauthner. "Whose Data Are They Anyway? Practical, Legal and Ethical Issues in Archiving Qualitative Research Data." *Sociology* 38 (2004): 139–52. Print.
- Poole, A. H. "How Has Your Science Data Grown? Digital Curation and the Human Factor: A Critical Literature Review." *Archival Science* 15.2 (2014): 101–39. Print. doi:[10.1007/s10502-014-9236-y](https://doi.org/10.1007/s10502-014-9236-y).
- Rasmussen, K. B., and G. Blank. "The Data Documentation Initiative: A Preservation Standard for Research." *Archival Science* 7 (2007): 55–71. Print. doi:[10.1007/s10502-006-9036-0](https://doi.org/10.1007/s10502-006-9036-0)
- Rouse, J. *Engaging Science: How to Understand its Practices Philosophically*. Ithaca and London: Cornell University Press, 1996.
- Sana, M., and A. Weinreb. "Insiders, Outsiders, and the Editing of Inconsistent Survey Data." *Sociological Methods & Research* 36.4 (May 2008): 515–41. Print.
- Savage, M. "Revisiting Classic Qualitative Studies." *Forum: Qualitative Social Research* 6.1 (2005): 31. Web. 28 May 2015. <<http://www.qualitative-research.net/index.php/fqs/article/view/502>>.

- . *Identities and Social Change in Britain Since 1940: The Politics of Method*. Oxford: Oxford University Press, 2010. Print.
- Schwartz, J. M., and T. Cook. “Archives, Records, and Power: The Making of Modern Memory.” *Archival Science* 2 (2002): 1–19. Print.
- Somers, M. “What is Sociology after the Historic Turn? Knowledge Cultures, Narrativity, and Historical Epistemologies.” *The Historic Turn in the Human Sciences*. Ed. T. J. McDonald; Ann Arbor, MI: University of Michigan, 1996. 53–90. Print.
- . *Genealogies of Citizenship: Markets, Statelessness, and the Right to Have Rights*. Cambridge: Cambridge University Press, 2008. Print.
- Stoler, A. L. “Colonial Archives and the Arts of Governance.” *Archival Science* 2, 1–2 (2002): 87–109. Print.
- Tamboukou, M. “Archival Research: Unravelling Space/Time/Matter Entanglements and fragments.” *Qualitative Research* 14 (2014): 617–633. Print.
- Thomson, R. “Possession: Research Practice in the Shadow of the Archive.” *The Craft of Knowledge: Experiences of Living with Data*. Ed. C. Smart, J. Hockey, and A. James; Palgrave Macmillan: Basingstoke, 2014. 39–55. Print.
- Trace, C. B. “What is Recorded is Never Simply ‘What Happened’: Record Keeping in Modern Organizational Culture.” *Archival Science* 2, 1–2 (2002): 137–59. Print.
- UK Data Archive. *Create & Manage Data: Anonymisation*. Web. 31 May 2015. <<http://www.data-archive.ac.uk/create-manage/consent-ethics/anonymisation?index=1>>.
- Whatley, P., and C. Brown. “Along and Against the Grain: The Philosophy of the Archive.” *Archival Science* 9 (2009): 127–31. Print.

APPENDIX 1. METADATA SHEET FORM CESSDA

The screenshot shows the metadata page for the dataset 'Election Funding of Finnish MPs 2007' on the Finnish Social Science Data Archive (CESSDA) website. The page is divided into two main sections: 'DESCRIPTION' and 'SUMMARY STATISTICS'.

DESCRIPTION:

- Dataset:** Election Funding of Finnish MPs 2007
- Variable Q14:** [q14] The amount of contributions received from party organisations (in euros)
- Literal Question:** The amount of contributions received from party organisations (in euros)
- Summary Statistics:**
 - Valid cases: 271
 - Standard deviation: 8096.337
 - Minimum: 0.0
 - Maximum: 57300.0
 - Mean: 4601.3

Yhteystiedot FSD | FSD contacts | fsi@uta.fi

Copyright © 2001-2006, NESO - All Rights Reserved - Contact: Newsletter, Support