

The Pragmatics of Metadata: From Practice to Concept to Practice

Deborah Maron

School of Information and Library Science
University of North Carolina at Chapel Hill
maron@live.unc.edu

ABSTRACT

In this paper, I describe building a conceptual model of metadata that is based on embodied, phenomenological experiences of practitioners from a wide variety of metadata-dependent fields. The pursuit will hopefully forge new horizons for metadata practice, pedagogy and scholarship while helping us see metadata as broader than what it is typically thought to be.

CCS Concepts

• Document Metadata → Digital Libraries and Archives • Information Retrieval

KEYWORDS

Metadata; Information Retrieval; Databases, Digital Libraries; Digital Archives, Cataloguing; Philosophy

1. INTRODUCTION

Few would deny that metadata permeates our world. Despite its ubiquity, however, metadata as a formal library world concept lacks a universal definition [5]. Defining metadata *practice* is similarly contentious. Colloquially (and sometimes in textual discourse) when the metadata community uses “metadata” to describe their professional work they typically mean 1) work in the area of Linked Data, or 2) XML encoding of records and or schema development using standards like MODS, METS, Dublin Core, etc. [5; 17; 1; 2]. This is curious, because if metadata practice is defined as the purposeful creation of data around information/knowledge objects for their retrieval and organization, then information retrieval, database design, interface design, and especially cataloguing/classification and digital archiving, are also metadata work. Separateness has a number of negative implications for metadata work, the most significant being that many of whom I consider to be “metadata professionals” are engaged in a Tower of Babel situation, unable to communicate with one another despite their common goal of wanting things to be well organized and retrievable. What if, however, the cataloguer, the digital archivist, and the algorithm writer understood each other better? We could more effectively develop metadata-dependent systems (e.g. digital libraries and retrieval layers on databases, or Natural language processing in digital archives) if workers could converse across their artificial boundaries. I propose that “step 1” in overcoming the communication hurdle is building a conceptual model of metadata practice premised on the belief that mutual understanding is promoted by talking to individuals actually doing the work, by looking at the problem of metadata literacy and practice as *embodied and phenomenological*. I assert that metadata as a critical, complex concept can only really be understood vis-a-vis the people who practice it; following that, a person’s metadata *literacy* is not quantifiable nor can it be determined by an individual asserting, in a survey instrument that ultimately disembodies them, that they possess an abstract set of skills. Rather, metadata literacy happens in specific, personal contexts that are defined by a person’s socio-economic, gendered, educational and work background, as

well as their orientation to the world. The way that a person performs metadata results from this mosaic of self as well as in the interactions with “materials specific to their situations”. In a study such as the one I am doing, the first person accounts—the embodied experiences that come into being through situated action— can be strung together to illuminate metadata as a large-scale social phenomenon. In sum: Metadata as a formal library world concept carries great potential. Redefining metadata as a formal library world concept is ultimately dependent on making those other practices, named above, legitimate members of the metadata practice community. Metadata then becomes an umbrella term under which many embodied and phenomenological practices are thematically, philosophically and intellectually bound in ways that make sense pedagogically and for the work that needs to be done.

2. RESEARCH QUESTIONS

I pursue three primary research questions in my dissertation work:

R1: What are the philosophical commitments/epistemologies of metadata practitioners?

R2: How do metadata workers characterize their work and themselves as workers?

R3: What do these things reveal about the concept of metadata?

3. PROBLEM DOMAIN AND STATE OF SOLUTIONS

Earlier I addressed the professional account of metadata. An interesting counterpoint to (or affirmation of) professional metadata can be found in scholarly literature. We know something about how scholars in domains like Knowledge Organization/Representation, Computer Science, Metadata Studies, Linguistics and Communication Studies regard this broadly construed idea of “metadata”, though these scholars might not explicitly call what they’re talking about “metadata”. Some literature in these domains is more speculative or prescriptive (e.g., Birger Hjørland suggests Critical Realism for viewing or performing our work, and Karen Wickett provides a set of axioms for connecting items to collection metadata), while other works are more descriptive or explanatory (e.g. Jens-Erik Mai’s critique of the Modern state of classification, or examples of Warrant exercised in a Fandom archive) [6; 16; 8; 4]. Others implicitly affirm the professional narrative by quantifying for and measuring skills of metadata workers in an abstract sense [10]. Others yet abstract or instrument the metadata work process itself (e.g. creating flow diagrams for how a cataloguer applies subject headings, or providing a Semiotic look at subject indexing) [12;9].

Authors’ philosophical commitments, or orientations to work and the world, are unmasked when analyzing scholarly literature. They can be, according to my research, roughly classified along a continuum of viewing the metadata person/product as universal, to

viewing the metadata person/product as socially dependent. Such views also have varying implications for meaning (e.g., is meaning “in” the metadata? Is it God-given? Do outsiders simply apply their own meaning in and around metadata?) However, as illustrative as this discussion may be, it is only so up to a point because it is still mostly abstracted from the actual people doing the work. As scholars and practitioners who write about metadata we have on the whole skipped “Step 1”: Similar to how Donald Schon in the *Reflective Practitioner* challenges a state of affairs in which scholars have a higher place in the world than practitioners and are dictating their practice without truly understanding it [13], I assert we should start from practitioner ground zero, analyzing the workers’ motives, beliefs and orientations to the world (i.e. epistemologies), their educational backgrounds, their workplace tendencies and mentorship opportunities, and how all these things shape metadata workers and the work they do. This knowledge of the real world, understood as an assemblage of personal and social philosophies, might then inform how scholarship is approached in “metadata” in the future (and of course, also help us to reconsider metadata as a formal library concept).

Metadata is understood vis-a-vis the practitioner. Therefore, rationality of metadata workers is especially consequential in reframing metadata. *Instrumental (means-ends) rationality* is the de-facto standard for “metadata thinking” and predominates in professional metadata [13]. This mindset around rigor, a reverence for numbers and “getting the job done” is challenged by reflexivity. Schon emphasizes reflexivity in action (i.e. being in conversation with the materials- be they physical, mental, or otherwise- that comprise your work situation) and honest, reflective conversations between work leaders, employees and colleagues as a means of advancing skilled professions. Reflexivity is similarly espoused in Feminist Epistemology, which asserts emotive, socio-economic and gendered qualities as positives for engaging with your work [15].

In 2005, a SIG call for papers implored classificationists to consider a “Descriptive Turn” in their scholarly work, focusing less on prescriptive angles and hard skills and more on descriptions of the embodied experience of classificatory endeavors [11]. Currently, metadata literacy is largely quantified in terms of things “that” we know rather than “how” we know, meaning that many scholarly articles depend on surveys and other tools measuring the skills possessed by, for example, librarians and archivists in a certain institution. These measured hard skills or technical benchmarks, which we see in everything from journal articles to job ads on AUTOCAT, reflect and dictate the supposed state of knowledge. I suspect that a conceptual model revealing embodied ability, beliefs and experience will trouble the state of metadata affairs in a fundamental way.

4. METHODOLOGY

Since the Enlightenment and arguably before it, natural scientists have regarded objectivity, numbers and measurement as the hallmarks of rigorous research, leaving little room for other, more interpretative techniques. Qualitative methods, namely in social science, emerged in reaction to this positivist or logical-pragmatic “Grand Narrative”. Kim writes:

Qualitative research informed by different interpretative paradigms uses words rather than numbers in its analyses and focuses on understanding human action through interpretation rather than prediction and control. [7]

I believe that metadata practice is embodied and cannot be properly characterized only in facts and figures; therefore, qualitative

techniques seem most appropriate to get at the heart of metadata as a formal library concept. Though there are many flavors of qualitative methodology, I will use the 3-part interview and Narrative Inquiry. The 3-part interview involves interviewing each subject three times and iteratively doing data analysis so that each subsequent interview can be more informed [14]. Narrative Inquiry is a rich methodology that combines the power of story with the interview technique. In Narrative Inquiry, the interviewee tells their story somewhat freely (with as much or as little scaffolding as you want to provide via your questions); such an approach allows you to understand the interviewee as a complex being. Narrative Inquiry has been shown to be a particularly effective method for interviewing practitioners, my population of interest:

Polkinghorne (1988) believes that working with stories holds significant promise for qualitative researchers because stories are particularly suited as a linguistic form in which human experience can be expressed. Narrative inquiry utilizes interdisciplinary interpretive lenses with theoretically, philosophically diverse approaches and methods, all revolving around the narratives and stories of research participants. [7]

Because I am interviewing people in pursuit of understanding and reconfiguring a concept, metadata, I will use also elements of the “concept” interview in my narrative interview, asking questions to probe interviewees’ experience with metadata in a conceptual sense. [3].

Narrative Inquiry entails numerous methods and philosophies, namely pragmatism (i.e. a belief that individual experiences tell us the truth), Heideggerian phenomenology (i.e. a belief that the subject (interviewee) and object (metadata) can only be understood together; the subject (interviewee) is “thrown”, or a part of, their particular experience), postmodernism (e.g. the belief that meaning is subjective, constructed and perspectival, dependent on the researcher), poststructuralism (e.g. one should be mindful of competing *discourses*, or presentations of belief, like between researcher and interviewee), and Socratic methods in interviewing (e.g. pushing back on a subject’s statement to see if there are any contradictions in what they say). [3;7]. I will also let a feminist epistemology determine the types of questions I ask during the interviews.

4.1 Positionality/Epistemology

Feminist epistemology guides my thinking because not only does it provide a useful framework for viewing your subjects, but it also takes into account that women have been doing the invisible work of library and information science for decades. This unappreciated work has shown up in numerous related fields, for not only is cataloguing a historically feminized profession, but “women computers” have also historically undertaken mathematical, problem solving endeavors at institutions such as NASA. Although women were (and are) integral for descriptive and computational work, their roles were (and perhaps still are) patriarchally dictated; that is, male ideals framed how women were viewed and employed. Feminist scholars later introduced the notion of the emotional worker to countervail positivism; likewise, concepts of intersectionality and standpoint epistemology, which consider gendered, socioeconomic and racial factors in understanding and analyzing workers male and female were also introduced. These viewpoints, and the feminine historicity of metadata, inform how I will conduct my study.

4.2 Data Analysis

At this point, I am considering different modalities of data analysis to be used together in bricolage [7].

At the conclusion of the three part interviews, I will create *profiles* [14] of each of my subjects to “objectively” and narratively piece together the most interesting or noteworthy things they said. From these, I can understand the stories of my participants and then further analyze the data, being confident that I have “read” the statements in a way that honors their integrity (as much as there can be a thing called “integrity” in this paradigm of thought).

I will then apply a critical commonsense understanding [3] to the subject matter during analysis, as the matter is complex, embodied, and is at its core commonsensical. I will ask: What precisely is the phenomenon of interest and what can different interpretative approaches tell us about the phenomenon? Is the phenomenon the person, metadata practice, metadata itself, or as I propose, all these things?

I will also employ dialectical (disagreement) and discourse analysis [3], which assumes neither genuineness nor inaccuracy in what someone says. Which discourses, or platforms of belief, predominate or appear in interviewing metadata people about their work? How will my ideas and interpretations of metadata (my discourse) jibe with what I find out?

Finally, I will consider the level of reflexivity between me and the subject, and between the subject and their work. To what extent are they in conversation with their materials? Are certain subsets of metadata work more or less reflexive than others?

5. HOW THIS CONTRIBUTES TO METADATA RESEARCH

Although scholars have created models showing how they believe we do subject indexing, cataloguing and other forms of classifying, we currently have no general conceptual model showcasing the epistemologies involved in metadata practice on the whole. Metadata as a thing is contentiously defined, as is practice, despite the fact that generally speaking we all want to better organize and retrieve information objects, whatever our particular “domain”. Metadata workers’ embodied metadata experiences tells us something about how we can unify divided practices that I expect are actually more alike than different on the fundamental level (or are different in trivial, methodological or semantic ways only). Instead of looking at an abstract set of hard skills, particular technologies or benchmarks or creating a process flow diagram of a narrow situation, I present a narrative and conceptual account of metadata as a social phenomenon that is perhaps more performative, person-oriented and nuanced than we had previously thought. My hope is that by understanding practitioners better in this embodied, highly personal fashion, we can develop a conceptual “lingua franca” that will benefit collaborative approaches to teaching and thinking about metadata in schools, workplaces and scholarship. In workplaces and in the classes we teach, the systems we create (databases, digital libraries, etc.) can potentially be improved with a renewed outlook on metadata as a formal library concept, because we will be remedying the unnecessary fractures in our field and bringing together people with different backgrounds, talents, and approaches to metadata problems. Choosing but one example, the information retrieval expert can, assuming a conceptual model is built that informs

praxis, contribute to digital archiving in ways that advance that domain, and likewise, she can learn from the digital archivist more about the library and archival mechanisms that assist in making organization and retrieval better. Metadata experts of different backgrounds and experiences might amend views they once held sacrosanct, acknowledging that other approaches could be useful. Harboring a different philosophical commitment from someone else can also be overcome; that is, in spite of one person being a “positivist” and another a “constructivist”, individuals can hopefully converse in the spirit of bettering metadata practice and systems reliant on metadata.

6. REFERENCES

- [1] Anon, 2017. American Library Association. Available at <http://www.ala.org>. Last accessed April 15, 2017.
- [2] Anon, 2005-2017. Upcoming Webinars. Available at <http://dublincore.org/resources/training/>. Last accessed April 15, 2017.
- [3] Brinkmann, S. and Kvale, S. 2014. *InterViews: Learning the Craft of Qualitative Research Interviewing*. SAGE Publications.
- [4] Bullard, J. 2017. Warrant as a Means to Study Classification System Design. *Journal of Documentation*, 75-90.
- [5] Coyle, K. 2004. Understanding Metadata and its Purpose. Available at <http://www.kcoyle.net/jal-31-2.html>. Last accessed April 15, 2017.
- [6] Hjørland, B. 2004. Arguments for Philosophical Realism in Library and Information Science. *Library Trends*, 488-506.
- [7] Kim, J. 2015. *Understanding Narrative Inquiry*. SAGE Publications.
- [8] Mai, J. 2011. The Modernity of Classification. *Journal of Documentation*, 710-730.
- [9] Mai, J. 2006. Semiotics and Indexing. *Journal Of Documentation*.
- [10] Park, J, Tasaka, Y and Maszaros, S. 2010. From Metadata Creation to Metadata Quality Control: Continuing Education Needs Among Cataloging and Metadata Professionals. *Journal of Education for Library and Information Science*, 158-176.
- [11] RebeccaG. 2005. Asis&t 2005: Day 1. The Dewey Blog. Available at <http://ddc.typepad.com/025431/2005/10/>. Last accessed April 15, 2017.
- [12] Sauperl, A. 1999. *Subject Determination during the Cataloguing Process*. Doctoral Thesis. University of North Carolina at Chapel Hill.
- [13] Schon, D. 1984. *The Reflective Practitioner*. Basic Books.
- [14] Seidman, I. 2012. *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*, Fourth Edition. Teacher’s College Press.
- [15] Sprague, J. 2005. *Feminist Methodologies for Critical Researchers: Bridging Differences*. Alta Mira Press.
- [16] Wickett, K. 2012. *Collection/Item Metadata Relationships*. Doctoral Thesis., University of Illinois at Urbana-Champaign.
- [17] Zeng, M. and Qin, J. 2016. *Metadata*. Neal-Schuman, Chicago, IL.