696

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

Metadata*

Matthew S. Mayernik

National Center for Atmospheric Research (NCAR), University Corporation for
Atmospheric Research (UCAR), P.O. Box 3000, Boulder, CO, USA, 80307-3000,
<mayernik@ucar.edu>

Matthew S. Mayernik is a project scientist and research data services specialist in the National Center for Atmospheric Research (NCAR) Library, based in Boulder, CO, USA. His work is focused on research and service development related to scientific data curation. His research interests include metadata practices and standards, data citation and identity, and social and institutional aspects of research data. He received his master's degree in library and information science and PhD from the UCLA Department of Information Studies.

Abstract: Metadata in various forms pervades our institutions, technologies, and daily lives. Metadata is a distinct focus of academic research and professional practice for many people within the library and information sciences (LIS). This article is an exploration of the concept of "metadata." It presents a high-level introduction to the topic with analysis of key research problems and practical challenges. The paper discusses varying understandings of what "metadata" means, the origin and evolution of metadata as an important topic within information and data fields, and the central characteristics of that which gets called "metadata." Metadata can be understood as both process and product and can result from both human effort and computational techniques. Given the central role metadata have in the establishment of knowledge, evidence, and truth, it is necessary for researchers and professionals within LIS to think critically about our metadata practices and systems.

1.0 Introduction

"The future province of metadata is grand" (Greenberg and Garoufallou 2013, 2). For many people within the library and information sciences (LIS), metadata is a distinct focus of academic research and professional practice. LIS is unique in putting such a lens on metadata as a matter of disciplinary emphasis, but as indicated by the epigraph above, the scope of the people and institutions who are interested in, or work with, metadata is indeed grand.

Outside of LIS, metadata has traditionally been a prototypical infrastructural phenomenon: essential yet mundane and ubiquitous yet often invisible (Borgman 2003; Edwards 2010; Pomerantz 2015). In the past decade, however, metadata has emerged as a critical topic in many contexts. Metadata became a topic of political and legal intrigue with the publishing of stories about the US National Security Agency eavesdropping on digital communications (Schneier 2014; Mayernik and Acker 2018) and the use and manipulation of metadata gathered by social media platforms (Acker 2018). Online streaming services for music, movies, and other forms of personal entertainment rely on metadata of various kinds to provide recommendations, personalization, and categorizations to their users (Madrigal 2014;

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

697

Maron and Carter 2017; Sisario 2019). And anything described as involving "big data," whether in the academic or business contexts, also inevitably involves "big metadata" (Greenberg 2017).

When viewed under a knowledge organization (KO) lens, metadata can be either (or both) something to be organized and something to use to achieve organization. As Richard Gartner (2016, 109) noted in a recent book on metadata, "Metadata is in many ways an attempt to develop a science for organizing ideas and so creating knowledge." Many of the long-term research questions in knowledge organization outlined by Gnoli (2008) have implications for metadata principles and practices, including "How can KO be adapted to local collection needs?," "How can KO deal with changes in knowledge?," "How can software and formats be improved to better serve KO needs?," and "Who should do KO?" Similarly, Hjørland's (2008) list of approaches to KO, including classification systems, facet analysis, information retrieval, bibliometric approaches, and the domain analytic approach, all often involve and/or manifest as some form of metadata.

1.1 Goals of the paper

This article is an exploration of the concept of "metadata." It presents a high-level introduction to the topic with analysis of key research problems and practical challenges. The theoretical view taken is of metadata as a sociotechnical phenomenon. Metadata, like data, comes from somewhere (Gitelman 2013). They have origins, histories, and journeys (Leonelli 2016). The intention of this paper is to discuss varying understandings of what "metadata" means, the origin and evolution of metadata as an important topic within information and data fields, and the central characteristics of that which gets called "metadata." The article discusses metadata as both process and product, illustrating how metadata is created and used within different kinds of contexts.

1.2 Topics beyond the scope of this paper

This article is not a "how to" document that will guide readers through particular metadata schemas or standards. There are many helpful book-length guides that provide in-depth instruction on many specific metadata languages, including those by Caplan (2003), Foulonneau and Riley (2008), Sicilia (2014), Zeng and Qin (2016), and Haynes (2017). The discussion in this article is complementary to those works, as well as to earlier overviews of metadata by Jane Greenberg (2005; 2009), by discussing metadata in the context of the people, technologies, and institutions with which they are connected. The article also does not focus on any particular technologies or intellectual domains as many

of the characteristics discussed below manifest across a range of technical infrastructures and institutions.

2.0 "Metadata" within library and information science

Although the exact origins of the term "metadata" have been recounted in different ways over the past couple of decades (c.f. Greenberg 2005; Giles 2011; Gartner 2016), the generally accepted view seems to be that the term originated in the late 1960s in the context of computer system design to refer to the use of one data element to describe or represent some characteristic of another data element. A search of the Web of Science citation indexes in December 2019 shows that usage of the term "metadata" first appears in 1982 with rare and idiosyncratic usage through the 1980s. It started to become a term of niche usage in the early 1990s in discussions of information management systems, geographical information systems, and database design. Aside from sporadic early use, the term entered the discourse of the library and information sciences in the mid-1990s, particularly in relation to the development of digital library systems and the emergence of the internet and the World Wide Web as major social forces.

The term "metadata" became widely used during the mid-1990s to refer to approaches to information description, management, and discovery that differed from conventional cataloging approaches using library-focused structure and content standards like Machine Readable Cataloging (MARC) and the *Anglo-American Cataloging Rules, Second Edition* (*AACR2*). The development of the Dublin Core metadata element set in 1995 exemplifies this turn toward "metadata" within the LIS communities, both in concept and terminology. The Dublin Core, so named because it was formulated in a workshop in Dublin, Ohio, in March of 1995, was explicitly motivated by a desire to develop a common approach to describing electronic resources that would enable better discovery and collection of resources on the web (Weibel, 1995; Sugimoto, Baker, and Weibel 2002). Keeping track of web pages and other digital information resources with traditional library cataloging practices proved to be difficult, despite the best efforts of library professionals, because of the malleable nature of internet-based materials.

Metadata subtag	Definition
Title	A name given to the resource
Creator	A person primarily responsible for making the content of the resource
Subject	The topic of the content of the resource
Description	An account of the content of the resource

698

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

Metadata subtag	Definition
Publisher	An organization or person responsible for making the resource available
Contributor	A person responsible for making contributions to the content of the resource
Date	The date that the resource was published or copyrighted
Type	The nature or genre of the content of the resource
Format	The physical or digital manifestation of the resource
Identifier	String or number used to uniquely identify the object, e.g., the object identifier (OID)
Source	A reference to a resource from which the present resource is derived
Language	The language of the intellectual content of the resource
Relation	A reference to a related resource
Coverage	The extent or scope of the content of the resource
Rights	Information about rights held in and over the resource for rights management

Table 1. Dublin Core metadata standard.

The Dublin Core element set included thirteen fields (later expanded to fifteen by including Description and Rights), as stated by Weibel (1995, n.p.) to be "the minimum number of metadata elements required to facilitate the discovery of document-like objects in a networked environment such as the Internet. The syntax was deliberately left unspecified as an implementation detail. The semantics of these elements was intended to be clear enough to be understood by a wide range of users."

Weibel's quote displays a couple of important points of debate that existed at the time (Lagoze, 1996) and continue to manifest in relation to metadata developments today.

– First, there is tradeoff in the extent of the metadata that is needed (e.g., minimal vs. comprehensive description) in relation the goals of the effort (document discovery in the case of the Dublin Core).
– Second, metadata initiatives often face challenges in defining an appropriate degree of standardization. In the case of the Dublin Core, it solved a certain kind of interoperability challenge by standardizing the names of the metadata elements but opened new interoperability challenges by not specifying the syntax of the information held by those elements.
– Third, in declaring that the Dublin Core was intended to be clear enough to be used by "a wide

range of users," the developers were explicitly going against prevailing approaches in which metadata standards and practices were targeted towards professional experts.

The Dublin Core thus exemplifies how the move toward "metadata" assumed and asserted that metadata descriptions for resources in the web environment would be created by a range of individuals, from expert to novice.

This last point particularly illustrates how early discussions explicitly centered on the ways that "metadata" existed as a counterpoint to conventional approaches to library cataloging (Greenberg 2005). Michael Gorman, editor of the *Anglo-American Cataloging Rules* (*AACR*) for many years, was a noted critic of the move toward metadata (Gorman 1999). In a later article simply titled "Metadata Dreaming," Gorman (2006) stated that the approach to metadata development and implementation exemplified by the Dublin Core was based on a failed utopian dream of a "third way" of description (with the bibliographic description approach and the free-text "Google search" approach being the other two ways). In a subsequent memoir, Gorman has called metadata "an inferior, unstandardized species of cataloging done by amateurs" (2011, 191) that is targeted towards the "philosopher's stone of bibliography—high-quality cataloging with no or little expense" (203). Gorman was perhaps one of the more visible and vocal critics but was by no means the only voice that argued that metadata projects that lose the more structured and complex approaches used in the library and archival community would struggle to be successful over the long term (Howarth 2005).

Such criticisms failed to slow the momentum of the "metadata" movement. The Dublin Core itself is the center of a dynamic metadata research and application community and is now a common reference point for many approaches to define "minimal metadata sets" for various purposes (Arakaki et al. 2018). It was also integrated into the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) as a way to facilitate harvesting of resources from diverse online sources (Van de Sompel et al 2004). The tensions noted above regarding metadata completeness, standardization, and consistency have not disappeared (c.f. Lagoze et al 2006; Urban 2014), but it is fair to say that the trends that motivated the development of the Dublin Core have held true. Namely, as most information and data systems have moved online and new types of internet-based information and communication technologies have emerged, the numbers of people, information types, and standards at play in the metadata space have increased correspondingly (Lagoze 2010).

The term "metadata" is now generally used in an expansive fashion to refer to descriptive and organizational schemes and practices broadly, regardless of whether they take place within information and data institutions or in other contexts. Like

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

699

its close relative "data," the term "metadata" has been used to function as a plural or a collective singular noun (Rosenberg 2013). "Metadata" is commonly used now as a blanket term for a range of practices, many of which pre-existed the term "metadata" itself, including library cataloging, archival description, and scientific data documentation, along with more recent phenomena, such as automatically generated information associated with digital images or social media streams (c.f. Pomerantz 2015; Gartner 2016; Haynes 2017).

3.0 Metadata definitions, conceptions, and relations

As the term "metadata" has spread, it has been defined and redefined in numerous ways. Many scholars and professionals have moved past the most common definition of metadata, the literal "data about data," to more nuanced and pragmatic discussions of requirements and functions.

3.1 Definitions

The following list provides a handful of definitions of metadata to illustrate how such definitions range from fairly specific to quite broad.

- Greenberg (2003, 1876)—"structured data about an object that supports functions associated with the designated object."
- Greenberg (2005, 20)—"data attributes that describe, provide context, indicate the quality, or document other object (or data) characteristics."
- Smiraglia (2005, 2)—"structured descriptors of information resources, designed to promote information retrieval."
- Gilliland (2008, n.p.)—"the sum total of what one can say about any information object at any level of aggregation."
- Pomerantz (2015, 26)—"Metadata is a statement about a potentially informative object."

Within particular application areas or academic communities, more targeted definitions appear as, for example, in the following set of definitions of metadata by experts in geo- and environmental science data.

- Michener, et al. (1997, 331)—"higher level information or instructions that describe the content, context, quality, structure, and accessibility of a specific data set."
- Fegraus et al. (2005, 159)—"the information that describes 'who, what, where, when, why, and how' an ecological data-set was collected."
- Danko (2012, 360)—"data that describes the information so that it will be useful and have value, be understandable, and enable collaboration."

- Gordon and Habermann (2018, 38)—"well-defined content in structured representations that make it easier to share and discover."

Jonathan Furner recently demonstrated how definitions of metadata also vary within standards established by ISO, the International Organization for Standardization (Furner 2020). Furner found that 96 ISO standards include definitions of "metadata," ranging from "data about data" to much more detailed definitions. Furner concludes that while one interpretation of these findings is that the ISO standards represent a problematic inconsistency in what "metadata" refers to within the information and data worlds, another interpretation is that these varying definitions represent community-centric interpretations of the "metadata" concept as appropriate for their applications. Furner (2020, E41) states:

> Perhaps it is actually more important for different domains, or even subdomains as represented by individual standards, to develop and to record their own particular definitions of terms, and thereby to make explicit the otherwise possibly overlooked differences in the ways in which the same terms are used in different contexts by different groups of specialists for different purposes.

One may argue that definitions in standards should take precedence over literature-based definitions. These two types of documents, however, tend to reach different audiences and, therefore, serve different purposes. Standards tend to be read and used by professionals while literature tends to be read by scholars and students. Therefore, neither source is more definitive. Rather, specific genres of documents (and specific individual documents) are more prominent within specific institutional situations. It is true that standards-based definitions are often developed by committees of individuals representing a variety of stakeholders. As Furner's study shows, however, differences between standards-making bodies and committees themselves are a source of variation among definitions of "metadata."

Definitions in both the research literature and standards, however, tend to focus on the use of metadata. This leads to the next topic, specifically, the ways that "metadata" have been categorized and conceptualized.

3.2 Categorizations and conceptions

If the notion of "metadata" has been defined in a variety of ways, it has been categorized and conceptualized in an even more diverse fashion. Categorizations of metadata reflect the different conceptions and motivations of the people who generate them and manifest in a variety of metadata ty-

700

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

pologies. Gilliland (2008), for example, in an overview article on metadata for library and information professionals, breaks the term "metadata" down into five types: administrative, descriptive, preservation, use, and technical. Other works, however, present different categorizations. A full explication and comparison of all of the categorizations that have been proposed is beyond the scope of this article, but looking across a selection of metadata-focused works (Greenberg 2001; Greenberg 2005; Lawrence et al 2009; Pomerantz 2015; Gartner 2016; Habermann 2018), the following categories appear at least once:

– access,
– administrative,
– archive,
– authentication,
– browse,
– character,
– descriptive,
– discovery,
– finding,
– identification,
– linking,
– preservation,
– provenance,
– relationships,
– rights,
– structural,
– technical,
– understanding,
– use.

Clearly some of these terms are related. But the variety of these categories indicates the generally broad understanding of what "metadata" might encompass.

Many of these categories represent particular tasks or actions that might be facilitated by metadata, such as authenticating, browsing, discovery/finding, preserving, or understanding information/data resources. This is perhaps the one commonality among these various definitions and categorizations of metadata: that metadata is created to be used, for some purpose(s), by people or computer applications. Karen Coyle (2010, 6) outlines how metadata is constructed, constructive, and actionable:

– Metadata is constructed: It is an artificial creation not found in nature.
– Metadata is constructive: It is created for a purpose, activity, or to solve a problem.
– Metadata is actionable: It is intended to be useful in some way.

Richard Gartner (2016, 4) provides a useful summation that encompasses this conceptualization of metadata as being designed and implemented for the purposes of particular uses:

The shape of metadata is designed by human beings for a particular purpose or to solve a particular problem, and the form it takes is indelibly stamped with its origins. There is nothing objective about metadata: it always makes a statement about the world, and this statement is subjective [in] what it includes, what it omits, where it draws its boundaries and in the terms it uses to describe it.

These characteristics of metadata hold across technologies, institutions, and decades. As one example, María Montenegro (2019) illustrates how the design of the Dublin Core metadata schema reflects the cultural assumptions of the people who were involved in its creation, particularly around notions of authorship and ownership of information resources. Information and knowledge that originate in other cultural contexts, such as within Indigenous communities, may not fit within the Dublin Core's framework. As Montenegro (2019, 737) notes:

Two DC [Dublin Core] elements in particular perpetuate colonial practices of exclusion. Specifically, the rights and Creator Fields conflict directly with Indigenous epistemologies and protocols defining the access, circulation and use of TK [Traditional Knowledge] … Both fields—rights and creator—are formed upon and replicate legal frameworks that have embedded relations of exclusion. The definition provided by DC for the rights element presumes that IP [Intellectual Property] laws are universal, however, legal regimes of IP and copyright are culturally specific and the types of rights they specify, by definition, exclude all types of Indigenous TK.

In another example, Fidler and Acker (2016) depict some of the decisions that were at play in the design of the protocols for information exchange that underlie the internet. The designers of the internet protocols engaged in a range of debates about what metadata needed to be associated with each "packet" of information that was transmitted over the network. Discussions took place about the importance of socket numbers, network addresses of the computers at each end of the transmissions, as well identifiers for specific computer processes that were to be invoked by the transmissions. Other pieces of metadata were discussed but ultimately not included in the protocol's requirements, including metadata related to the specific users who were doing the transmissions. These discussions were targeted toward

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

701

particular purposes, ranging from the technical functionalities that were desired to the need to potentially gather information for billing users for their usage of the network.

Such metadata can be found in the design of any networked information system; indeed, such systems cannot function without internal metadata that support the networks' communications and functions (Mayernik and Acker 2018). This makes problematic the notion of metadata (or data) as being "exhaust" within technical systems. This "exhaust" metaphor has become increasingly common in discussions of metadata within digital systems (c.f. Mayer-Schonberger and Cukier 2013; Schneier 2015; Edwards 2017). Pomerantz (2015, 126) for example, states that "Up to this point, 'metadata' has meant data that was created deliberately; data exhaust, on the contrary, is produced incidentally as a result of doing other things." As we see from the Gartner quote above and the example from Fidler and Acker, any metadata created automatically within digital/networked information systems is a designed feature. There is nothing incidental about its creation. Using such metaphors as "exhaust," "smog," or "waste" when talking about metadata "implies that these traces are inevitable, a by-product of human and technical activities that cannot be avoided, and once produced are out of human control" (Mayernik and Acker 2018, 178). These metaphors serve to obscure understanding of metadata rather than illuminate it.

3.3 Relation to other concepts

This section provides brief overviews of how the concept of "metadata" relates to other important concepts within the library and information sciences.

3.3.1 Data

If we take the literal "data about data" definition of metadata, then it is straightforward to say that "metadata" is simply a sub-category of "data." This is useful in that it allows us to characterize metadata as having certain properties that prior analyses have denoted for things classed as "data." Two definitions of "data" are provided here as illustrations. Furner (2017, 66) defines data as "concrete instantiations of symbolic representations of descriptive propositions, informed by empirical observation, about the quantitative and qualitative properties of real-world phenomena." Hjørland (2018) builds on Kaase (2001, 3251) to provide a more streamlined and generalized formulation: "Data are information on properties of units of analysis." Both definitions note that "data" refers to entities that represent or contain information ("symbolic representations of descriptive propositions" in Furner's terms) about other entities, whether "quantitative and qualitative properties of real-

world phenomena" or "units of analysis." Put more simply, Furner refers to data as "representational concreta," that is, something concrete (i.e., materially manifesting via some real-world entity) that represents something else.

In this sense, it is straightforward to consider "metadata" to be a kind of "data." The idea that "metadata" refers to "a statement about a potentially informative object" (Pomerantz 2015, 26), or other more specific definitions given above, fits well within the broad category of "data" as "representational concreta."

Simply considering metadata to be a sub-class of data in this way is unsatisfactory, however, in that it does not provide any insight into why one might call a given entity "metadata" instead of "data." Common distinctions within KO contexts, where metadata such as classifications or subject terms are greatly distilled representations or descriptions of informational resources, do not hold in some technical systems where the "metadata" stored by the system can be much bigger than the "data" (Klensin 1995; Brunton 2016).

Additionally, many pieces of information that are conventionally called "metadata" are in fact used by researchers and professionals as evidence to make particular claims. Think of the field of bibliometrics, for example, or the recent discussion of "bibliographic data science" by Lahti, et al (2019). Using something as evidence for specific claims is a key definition of "data" in the context of scholarly research, according to multiple recent scholars (Borgman 2015; Leonelli 2016). Mayernik (2019) argued that the distinction between data and metadata may be related to what is foregrounded and what is backgrounded in the context of a knowledge claim. In particular, metadata, "however instantiated in local situated activities of scientific research, are central to enabling something to serve an evidentiary role, that is, to *serve as* data. In particular, if data are entities used as evidence, then metadata are the processes and products that enable those entities to be accountable as evidence" (Mayernik 2019, 734-735, italics in original).

In sum, calling something "metadata" as opposed to "data" is a culturally contextual classification that rarely has a self-evident rationale (Boellstorff 2013). Data and metadata are often designated in contradistinction to each other, depending on the specific situations of origin and use (Borgman, Wallis, and Mayernik 2012; Mayernik and Acker 2018).

3.3.2 Document

The notion of a document is central to library and information science. Library and information work, including knowledge organization, centrally involves the creation, processing, and organization of documents. Library and information science scholars have thus developed sophisti-

702

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

cated understandings of what it means to call something a "document." Michael Buckland (1997; 2014) outlines how particular entities can be "made as," "made into," and "considered as" documents. These three views, which are progressively more inclusive, reflect how: 1) particular things may be deliberately designed to serve documentary purposes ("made as documents"); 2) human artifacts may be used as documentary resources even if that was not their original purpose ("made into documents"); or, 3) naturally occurring objects such as rocks or animals may be used for documentary purposes ("considered as documents"). In these senses, almost any object could be used as a document depending on their evidentiary value in particular circumstances. Being a "document" is, therefore, a role that particular things play, rather than an inherent property of those things.

Furner (2016, 303) argues that all datasets are documents, stating that "the dataset is a species of document." Thus, if metadata are a special kind of data, as noted above, then metadata likewise exist as documents not abstract concepts or information that exists without material form. As such, metadata can be analyzed via the same conceptual apparatus as documents. See Buckland (2018) for a recent overview on this topic.

If metadata is a sub-species of both data and document, it might be worth asking the question about the relationship among all three concepts. In other words, using Venn diagram terminology, one possible view is that the three terms are completely hierarchical, with "metadata" being completely encircled by "data," which is in turn completely encircled by "document." An alternate view is that "data" and "metadata" are partially overlapping circles within the larger "document" set. This latter view is a better fit with the discussion in the previous section of the culturally contextual ways in which data and metadata are distinguished. In other words, documents can be data, metadata, both, or neither depending on their usage as such in particular situations. Stated more concretely in relation to Buckland's conception of "documents" as being roles rather than properties, being metadata is a role that some documents (or types of documents) have in particular circumstances (Renear and Wickett 2010).

3.3.3 Context

A couple of the definitions of metadata presented in section 3.1 referred to metadata as describing or providing context for informational/data resources. "Context" is itself a potentially slippery concept, generally referring to the setting or situation in which an action or event takes place and the factors that influence the action or event as it happens. Contexts can be important in how metadata is designed or in how it is interpreted (Wickett 2015). Dervin (1997, 14)

notes that context is typically conceptualized, "usually implicitly, as a kind of container in which the phenomenon resides." Talja, Keso, and Pietilainen (1999, 754) approach context from a metatheoretical viewpoint, saying "context is the site where a phenomenon is constituted as an object to us." They describe context as the "crossroads between researcher and data." Dourish (2004) describes how what is usually referred to as "context" can be better conceived as being rooted in "practices." Shifting from "context" to "practices" allows us to focus on the "engaged action around artefacts and information that make those artefacts meaningful and relevant to people" (Dourish 2004, 26). Using this view, "context" exists both; 1) independently from the situated actions; and 2) co-produced by people via their situated actions. Metadata thus serve to create the context around information/data resources as much as they serve to describe that context.

As more metadata is produced automatically via computing systems, context is something that "must be reckoned in both architectural and institutional terms" (Agre 2001a, 194). In other words, "context," in the context of computing systems, includes considerations of both the operations of the computing hardware and software—from bit-level to infrastructure-level—as well as considerations of institutional settings in which those computing systems were designed, created, and operated. Metadata associated with digital objects may be designed to reflect different parts of these details, depending on the application or situation.

4.0 Characteristics of metadata

As the diversity of the definitions, functions, and roles given above illustrates, metadata is not a definite and singular concept. Rather, it is a fluid, multiple, and fractional concept (Law 2004). Metadata is "fluid" in that file naming conventions, catalog records, data descriptions in repositories, user tags on YouTube, notes in personal Excel spreadsheets, email headers, and HTML tags can all be called "metadata." Metadata, as a concept, is also characterized by "multiplicity" in that it is enacted differently in different social settings and situations, from Dublin Core records created by information professionals to descriptions in lab notebooks created by scientists to document their data. Despite this diversity, some characteristics and points of debate are common across metadata of different kinds. This section discusses some of the central characteristics of that which gets called "metadata."

4.1 Structured vs. unstructured

A primary point of distinction in some discussions of metadata is between structured and unstructured information. A number of the definitions quoted in Section 3.1 explicitly

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

703

call out metadata as being "structured." Many structures for metadata have been formalized into standards, ranging from general purpose metadata standards such as the Dublin Core to discipline-specific standards for particular kinds of resources, e.g., geospatial information (Danko 2012; Brodeur et al 2019). Standardized schemas and structures facilitate the interoperability of metadata between systems and applications (Zeng 2019). Metadata standards are commonly organized around a set of elements (such as "title," "author," "date") that manifest as computer-readable documents in one of an alphabet-soup set of formats and mark-up languages, such as MARC, XML, JSON, and YAML.

Structured metadata can be differentiated from other forms of unstructured metadata, which might also be called "documentation" (Habermann 2018). Unstructured metadata could include any range of traces and practices that achieve some or most of the same goals as structured metadata, namely to create documentation, descriptions, and annotations for the purposes of managing, discovering, accessing, using, sharing, and preserving informational/data resources. As one example, in the context of data archives, it is common to include one or more narrative documents that describe various aspects of the data in more detail than is possible through standardized metadata structures.

It is important to keep in mind that structured and unstructured metadata can be hard to fully disentangle. Metadata standards commonly include a mix of controlled and uncontrolled elements. Controlled elements may require the information therein to conform to a specified syntax (e.g., "year-month-day" syntax in a date field) or to be chosen from a pre-determined set of values (i.e., controlled vocabularies). Uncontrolled fields, on the other hand, may allow any value to be present. Thus, even within highly structured metadata standards, there can be significant amounts of unstructured metadata. This characteristic can challenge attempts to aggregate or discover metadata, even if it is all structured according to a common standard (Arms et al 2002).

4.2. Metadata-as-product and metadata-as-process

The use of standards to create structured metadata results in what can be characterized as "metadata products." Edwards et al (2011), in a discussion of metadata in the context of scientific research, describe how metadata products almost always involve corresponding "metadata processes," namely, practices that help people overcome or bypass frictions that occur in the creation and use of metadata. Edwards et al (2011, 684) states:

> Well-codified metadata products increase the precision with which a dataset can be fitted to purposes for which it was not originally intended, or can be reused

by people who did not participate in creating it. At the same time, ephemeral, incomplete, ad hoc metadata processes act as lubricants in disjointed, imprecise scientific communication. This latter category of metadata frequently appears alone, in the case of datasets for which no metadata products exist, but it also frequently appears in the actual use of metadata products.

Some of the examples provided by Edwards et al (2011) and other related works (Mayernik, 2019) discuss how "metadata processes" effectively serve to facilitate data discovery, sharing, and use in situations where standardized "metadata products" have not or cannot be created due to time constraints or the lack of expertise available. As noted in the last sentence of this quote, however, metadata processes are also important in situations where standardized metadata products are being created. As one example, starting in 1988 and extending into the 2010s, the US Library of Congress published a set of "rule interpretations" for use by catalogers within libraries across the world who were creating catalog records via the *Anglo-American Catalog Rules, 2ⁿᵈ Edition* (*AACR2*). *AACR2* provided hundreds of rules for cataloging library resources of all kinds. Applying these rules when cataloging particular items, however, involved interpretive decision-making regarding their fit to the details of the item in hand. The Library of Congress rule interpretations gave catalogers more detailed guidance on how to apply cataloging rules than was contained in the *AACR2* cataloging code itself. These rule interpretations covered common cases, such as how to enter author names when there was more than one author of a resource and rare cases, such as how to designate authorship for a conference proceedings, where no individuals were named as authors or editors. As Barbara Tillett, former head of the Library of Congress's cataloging division, noted (Tillett 2003, 113):

> These rule interpretations lead to greater consistency in applying the rules, which is important for a very large institution and for its partners who help create compatible bibliographic and authority records. These guidelines are not appropriate for a cataloging code, but are needed for training and daily guidance to catalogers seeking to provide bibliographic description and access in a consistent way.

The Library of Congress Rule Interpretations (LCRI) were thus a kind of metadata process that facilitated the creation of more standardized metadata products. The implementation of any metadata standard is tied up in local interpretations and processes (Park and Maszaros 2009). This interpretive flexibility is a characteristic of every metadata standard or schema (Feinberg 2017). Looking closely at the pro-

704

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

duction of other metadata products would likely show similar couplings with attendant metadata processes.

4.3 Metadata and description

In a recent work, Michael Buckland (2017, 113) states that the "first and original use of metadata is to describe documents." It is thus important to discuss briefly what "descriptions" are and what characteristics they pass on to metadata. The word "description," like other similar words such as communication, illustration, and, yes, information, can be used to discuss both things and activities. When talking about metadata, descriptions are most commonly discussed as things, e.g., descriptions of library resources, archival materials, or data sets that are held in information systems and made accessible through catalogs. Decades of sociological research, however, has focused on description as an activity. This literature cannot be fully detailed here, but it provides important insight into how descriptions created and used as metadata should be understood.

Descriptions, whether verbal or written, are "only more or less reliable by virtue of their being treated that way for the practical purposes at hand" (Woolgar 1981, 509). In this sense, metadata encompasses negotiated shared meanings. Metadata is typically created with the expectation that readers or users of the descriptions will have knowledge of how to read and interpret them. As Heritage (1984, 150-151, italics in original) states, however, "no description is strictly *compelled* by the state of affairs it describes. Any description is thus inherently *selective* in relation to the state of affairs it depicts ... [C]hoices which underlie any description ... are all sources of clues concerning how the description is to be interpreted." This characteristic, that metadata is inevitably selective, relates to the point in Section 3.2 above about metadata being created for specific purposes. Analysis of the metadata creation process should thus view metadata description, whether catalog records, classifications, labels, or technical traces, as a kind of action situated in social settings. In fact, as noted in the last sentence of the Heritage quote above, the metadata that does exist in some information system or social setting can itself be studied as a way to gather insight into the priorities, expectations, and accountabilities that exist in relation to those systems or settings (Mayernik 2019).

4.4 Search, discovery, and understanding

Beyond description, Buckland (2017, 118) notes that an additional use of metadata is to enable searching. Metadata can be used to provide structures that support consistent search and discovery of information across broad ranges of documents. Metadata can also potentially enable distinctions to be made among similar kinds of documents or re-

sources. A search in a library catalog for "Hamlet" or a search in a scientific data catalog for "climate data" can result in hundreds or thousands of relevant results. Metadata that is useful for search and discovery may not be useful in distinguishing the differences among such large numbers of results. Users will likely need additional metadata that allows them to understand the resources, not just discover them (Habermann 2018). Providing metadata for understanding is certainly a role of the descriptions noted in the previous section. For example, annotated bibliographies of the various editions of Hamlet (Bevington 2019) and comparative guides for climate data (Schneider et al 2013) exist specifically to go beyond search and discovery to enable understanding. Metadata is not the only way to move beyond searching to understanding. Interface design and better search capabilities also have an impact (Marchionini 2006). But additional and novel metadata kinds and structures are central to this goal.

4.5 Relationships

One critical characteristic of metadata is that they are often the carrier of information about relationships within, among, and between informational/data resources. Many information and data systems manage and leverage relationships of a variety of kinds, including relationships among vocabulary terms and content structures (Bean and Green 2001) and relationships between documents and networks of documents (Mayernik 2018). Research in knowledge organization centers heavily on how to understand and represent relationships, both of a conceptual and a documentary nature (Green 2008; Szostak 2012) and has defined canonical types of relationships that obtain in the information arena, including hierarchical, associative, and equivalence relationships (Bean and Green 2001).

Yet, this aspect of metadata is often underappreciated. As Geoffrey Bowker (2016, n.p.) noted, "we don't build our archives around relationships, we build them around things (if there is one fundamental flaw in our generic archival practices, it is this)." Gary Marchionini (2012), in an acceptance speech upon winning the "Award of Merit" from the Association for Information Science and Technology (ASIS&T) in 2011, suggested that "information science is in search of a theory of relationships," (20) and stated that the community would benefit from paying attention "to the nature of relations in general rather than only identifying specific new relations" (21).

Rebecca Green (2008) discussed a variety of ways in which relationships manifest in knowledge organization systems. Relationships might be expressed via classification systems, vocabularies and thesauri, subject headings, or via specific relationship-focused elements in bibliographic records. Recent developments in the semantic web centrally in-

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

705

volve the precise specification of relationships between entities (Allison-Cassin 2012; Dunsire, Hillmann, and Phipps 2012). All of these manifest as metadata in some kind of document and/or information system. When relationship metadata exists in defined and structured form, they can be leveraged within information systems to enable information discovery and understanding, as well as to allow properties of one item to be transferred or inferred to another (Wickett 2018). When relationship metadata exists as unstructured information, e.g., as components of narrative metadata, they can enable keyword-based searching or be used by users to better understand the item(s) in hand.

5.0 Where does metadata come from?

As noted in the introduction, metadata comes from somewhere (Gitelman 2013). The social settings in which metadata are created have a large impact on what form metadata takes and on who or what creates metadata. Metadata can be created through both automated and manual processes. Both of these methods present challenges. This section outlines different people and technologies that have roles and responsibilities related to metadata creation.

5.1 Professional metadata creators

In libraries and archives, the creation of metadata is an institutionalized task. Catalogers, archivists, and professionals with titles like "metadata librarian" (Han and Hswe 2010) are assigned responsibility for creating metadata. Within these kinds of institutions, metadata work is also frequently conducted by paraprofessionals who have knowledge, experience, and training with the relevant systems and standards (Moulaison Sandy and Dykas 2016). Metadata creation is also a key responsibility for people working as data managers within data repositories (Palmer et al 2014; Rasmussen 2014).

Researchers and professionals from the library and information sciences (LIS) often approach their work through developing and applying defined sets of principles. Principles are discussed at professional meetings and in the literature, debated in standard-writing committees, and taught as part of professional education curricula. Principles offer directives for how information systems and the languages they use should be designed (Svenonius 2000). Principles depict how things should be or would be in optimal circumstances (Gnoli 2012).

The articulation of principles has been a central activity (and point of debate) within the library cataloging community for decades. Cataloging codes since the 1960s have been based in community-accepted principles, starting with the "Paris Principles," which resulted from an international meeting held in Paris in 1961 (International Conference on

Cataloguing Principles 1971). In the mid-1990s, when new cataloging code revisions were being debated, no fewer than three international conferences were held that focused either in whole or in part on the fundamental principles that should underlie the next code (Weihs 1998; Schottlaender 1998; Harkness Connell and Maxwell 2000). Individual contributions to these conferences debated the implementation of principles in past codes and presented new principles for a variety of specific issues, such as principles for cataloging relationships between resources and principles for cataloging serial materials. The cataloging code that resulted from these debates, titled *Resource Description and Access*, includes a statement of principles in the introductory chapter and notes at the beginning of each subsequent chapter how each section of rules relates to the stated principles (JSC 2014). The development of archival practices and institutions since the nineteenth century has been likewise driven by principle-based approaches (Gilliland 2014) as was the development of the Dublin Core Metadata Schema in the 1990s and early 2000s (Weibel 1995; Arakaki et al. 2018). Information and data professionals are far from having a monopoly on metadata creation, however, especially if the scope of what "metadata" entails is taken broadly.

5.2 Automatic metadata generation

As noted above, digital systems are inherently dependent on metadata that is automatically created for a variety of purposes (Mayernik and Acker 2018). The more structured the digital workflow, the easier it is to automate the creation of metadata, for example, to record provenance information about how information or data have been derived or changed over time. Beyond the use of automation to generate technical metadata, however, automation can also be applied to generate descriptive or topical metadata. Jane Greenberg (2004) described how automated metadata creation techniques typically follow one of two approaches, extraction or harvesting. In metadata extraction, "an algorithm automatically extracts metadata from a resource's content" (Greenberg 2004, 62). Common applications of the extraction approach include automatic abstract generation for publications and summary displays of web pages given by web search systems. Metadata harvesting, on the other hand, involves compiling metadata automatically from distributed resources, such as collecting standardized metadata from metadata feeds or web site HTML. As Greenberg (2004, 63) notes, "the 'harvesting process' relies on the metadata produced by humans or by full or semi-automatic processes supported by software."

Automated metadata extraction and harvesting methods are most robust for text-based documents, for time- and geotagging of digital photos, and the like. But new technologies and techniques for extracting information from audio

706

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

recordings, video, and images also have promise for the purposes of metadata creation (Riley 2017). Facial recognition software could be used, for example, to create metadata about the people that are shown in digital video or image collections maintained by libraries and archives. Given the explosion of digital media and the growth of digital archives, these kinds of techniques may be the only tractable way for such metadata to be produced (see for example Mühling et al 2019). The use of facial recognition and other similar algorithmic metadata extraction techniques must be coupled, however, with strong awareness of the notable ethical implications that arise when creating information about people without their awareness (Agre 2001b; Seeman 2012; Crawford 2019; Padilla 2019).

5.3 Metadata creation in everyday life

Outside of information/data institutions and structured technological workflows, metadata creation can take on various forms with many opportunities and challenges. Many of the computational techniques noted in the previous section are either in nascent form or are not effective when applied to unstructured or very diverse informational resources. They may also require specific technical skills to implement. In day-to-day life, people may create metadata for information and data in the context of work or home settings. People may create metadata via folder structures and file names for personal images, or create one-off notes documents for particular tasks or resources. All of these are acts of metadata creation in the general sense.

Internet-based tools for sharing photos, videos, and other kinds of information commonly enable users to add metadata as tags to text or objects. Tags are also common within social media systems, where users of Twitter, Facebook, or Instagram include hashtags to connect their posts with other discussions within the platforms, such as #WomensHistoryMonth, #EmployeeAppreciationDay, or #data. Library and information science researchers have studied how the aggregation of such tags can create folksonomies that reflect the vocabularies and language usage of everyday people in contrast to the structured and pre-determined taxonomies created and used by information professionals. Folksonomies have been studied and implemented as ways to bridge between expert and non-expert vocabularies (Cairns 2013) and potentially to feed into the creation of formal taxonomies or ontologies (Gil et al 2017). Such "crowdsourcing" of metadata has benefits and drawbacks. Enabling users to add metadata via their own terminologies tend to better support information browsing than searching (Sinclair and Cardew-Hall 2007) but can be very effective at accommodating and celebrating multiple voices and perspectives on the resources being described (Srinivasan et al 2009).

Metadata created for everyday tasks should be expected to have different characteristics than metadata created by professionals for institutional purposes. Such metadata tends to be idiosyncratic, varying in content and structure from individual to individual, and from situation to situation for the same individual. The aforementioned discussion of principles for metadata do not apply. As Chamberlain and Crabtree (2016, 569) note in a study of how metadata is created and used in the context of personal music collections:

> Relevance is a key factor to understanding the nature of metadata, what is relevant in one context may rapidly change as different artifacts, reasons and results are employed indifferent emerging contexts – metadata is not always a static "entity," in many respects it consists of different physical modalities, relates to people (trusted) and has different perceived and actual temporal qualities. Our fieldwork shows that the emergence and use of metadata is both part of, and yet can be separated from the workflow.

Creating metadata in everyday situations is often a task that has implicit or explicit moral implications (Vertesi et al 2016). People feel morally responsible for keeping track of important documents and information, such as family pictures, yearly tax documents, or vaccination records for children. These types of documents may exist in a variety of digital and analog formats and in multiple technical environments (e.g., email, personal computers, mobile phones). Creating metadata and organization systems for these kinds of things can thus be a source of emotional and interpersonal stress. People thus make context-specific choices about what metadata to create and in what forms and systems.

5.4 Metadata collaborations

Metadata creation often involves collaborations between people with varied expertise and knowledge. People with special domain or disciplinary knowledge may seek out metadata experts for help in the creation of specialized collections of resources on a particular topic. Or libraries and archives may bring in experts on a particular topic to provide consulting or specialized knowledge for particular collections. Additionally, as more libraries and archives are collecting born-digital resources, such as digital datasets, software packages, and other kinds of materials, they often have multi-step workflows where the contributor of the resource is required to create metadata for their asset(s), with professional librarians or metadata experts providing review and quality control for the resulting metadata.

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

707

Such collaborations can be challenging. People with different perspectives will bring different expectations about technologies, workflows, and outcomes (Khoo and Hall 2013). The time, energy, and attention involved in creating, collecting, assembling, checking, and/or understanding metadata can be significant, particularly for people without experience in creating structured or non-structured metadata. For example, scholarly research data repositories commonly experience difficulty in getting data creators to create metadata, and the metadata that is created can be of minimal quality (Jones et al 2006; Bhandary et al 2018). In some cases, researchers will refrain from sharing data, because it takes too much effort to produce the data and associated documentation necessary for its use (Tenopir et al 2011).

The benefits of such collaborations, on the other hand, center on being able to take advantage of the different sets of expertise that different people may bring. For example, disciplinary topic experts have firsthand knowledge about how resources related to their areas of expertise are created and used, and thus can provide useful insight into what metadata should be created, and how the metadata may be optimally structured to support use and re-use (White 2010). In the best circumstances, metadata experts and non-experts work together (Gazan 2003). Information and data professionals may serve as intermediaries (Mayernik 2016) to support the optimal usage of the applicable metadata standards, vocabularies, while topical experts provide relevant metadata content and guidance on usage.

6.0 Metadata futures—conclusions and research questions

This paper begins with a quote stating that the future of metadata is grand. The subsequent sections illustrated how metadata in various forms pervades our institutions, technologies, and daily lives. The ongoing digitization of our societies is, if anything, accelerating this trend. As noted by Richard Gartner (2016, 96):

> The growth of the digital seems to need more metadata not less. Google and its peers make it possible to discover new material in ways which we could not have conceived of before but they need to be complemented by human thought and the metadata by which it is focussed.

Going forward, researchers and professionals will continue to grapple with long-lasting challenges related to metadata creation and use, including questions about how to negotiate cost/benefit trade-offs between structured and unstructured metadata and between human and machine generated metadata. But it is also clear that each new generation of information and data technology produce and require differ-

ent kinds of metadata than systems that existed before. Social media technologies demonstrate a trend toward what Ronald E. Day (2019) has called "a posteriori" metadata generation. Day uses this term to contrast with "a priori" metadata generation as in library cataloging and classification, semantic web ontologies, and scientific data catalogs, where the generation of metadata tends to come at the beginning of the life cycle of information/data use. In systems based on "a posteriori" metadata, the metadata that get generated, stored, and used within the systems are less focused on the properties of the entities within those systems, instead focusing more on what those entities do (or what is done to them by other entities). Twitter, for example, may collect metadata about particular Twitter users, but they monetize metadata that reflects what those users do. Likewise, Twitter generates metadata about each post, but they monetize metadata that reflects how those posts travel (how many likes, retweets, and replies are generated), and the social networks connected to those interactions.

Information and data researchers and professionals who have expertise in the "a priori" metadata approach have much to consider in this trend toward "a posteriori" metadata. As Day (2019, 138, italics in original) suggests, "*A priori* categories, such as those that result from classification structures, can be heuristics for investigating entities, but they are only that." What does knowledge organization consist of if the value and meaning of particular information and data resources is "based on statistical calculations of the use and the relations of data" (141) instead of a priori decisions about subjects, classes, and categorizations based on inherent properties of those resources? Open research questions on this point relate to the relative value of both approaches to metadata creation in relation to the goals of the metadata being produced. In section 3.2, I listed nineteen different types of metadata that have been identified in prior literature (administrative, descriptive, discovery, preservation, provenance, technical, etc.). Some of these metadata types are obviously conducive to the "a posteriori" approach, such as provenance metadata, but with other types, such as descriptive and discovery metadata, the relative value of the "a priori" vs "a posteriori" approach to metadata creation is still an open question.

Day's reflections likewise provide important considerations for the future of metadata and knowledge organization in relation to evidence and evidentiariness in the digital age. In a time of "fake news" and "deepfakes," that which is considered knowledge is tied to the kinds of evidence that exist to buttress knowledge claims and to the ways that that evidence is marshalled. The use of documents as evidence has always been tied up with assessments of the authenticity of the documents involved. Authenticity assessments in the digital age rely on metadata to provide (and create) context and accountability for evidence. When the metadata neces-

708

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

sary for these authenticity assessments is itself subject to manipulation and/or capitalization, the grounds for using particular documents or traces as evidence can be eroded (Acker 2018). Metadata can serve as a kind of capital (Greenberg et al 2014), that is, economic assets that take time, energy, and money to create, compile, and leverage and provide value to those who control and use them. Thus, open research questions exist about the epistemology of metadata, namely, how metadata comes to exist, how people and technologies can learn about their origins, how metadata relates to economic and political interests, and how such interests can also be known and made transparent, potentially through the very metadata that is under question. What is the role of metadata in supporting evidentiary claims in the digital age? Tracing and documenting of relationships is a key challenge in relation to this question. In digital systems and applications, metadata moves around, being transmitted, transformed, aggregated, or pulled apart based on the needs and interests of various stakeholders. Sometimes such metadata journeys are pre-defined, but they can also occur opportunistically (Leonelli 2016). Documenting these journeys involves a kind of reflexivity, where metadata, its origins, histories, and journeys, must be documented through (and as) metadata. Open questions exist about how metadata journeys and relationships can reflect the "fluid" nature of our societies and cultural practices (Srinivasan and Huang 2005).

As the kinds of metadata that are created and the sites of metadata creation continue to expand, these challenges of authenticity of documentary evidence and, therefore, knowledge generation and organization will continue to grow. As stated by Day (2019, 49), "truth is not transcendental, but rather is revealed by performative practices. Truth becomes evident at indexical points of revelation made possible through technologies of informational inscription." Metadata thus have a central role in the establishment of knowledge, evidence, and truth. Going forward, thinking critically about our metadata practices and systems will be essential for the ongoing evolution of our information-centric technologies and societies (Feinberg 2018).

References

Acker, Amelia. 2018. *Data Craft: The Manipulation of Social Media Metadata*. New York: Data & Society Research Institute.

Agre, Philip E. 2001a. "Changing Places: Contexts of Awareness in Computing." *Human-Computer Interaction* 16, no. 2-4: 177-92. doi:10.1207/S15327051HCI16 234_04

Agre, Philip E. 2001b. "Your Face is Not a Bar Code: Arguments Against Automatic Face Recognition in Public Places." *Whole Earth* 106: 74-7.

Allison-Cassin, Stacy. 2012. "The Possibility of the Infinite Library: Exploring the Conceptual Boundaries of Works and Texts of Bibliographic Description." *Journal of Library Metadata* 12, no. 2-3: 294-309. doi:10.1080/19386389.2012.700606

Arakaki, Felipe Augusto, Vesu Alves, Rachel Cristina, Amorim da Costa Santos and Placida Leopoldina Ventura. 2018. "Dublin Core: State of Art (1995 to 2015)". *Informacao & sociedade-estudos* 28, no. 2: 7-20.

Arms, William Y., Diane Hillmann, Carl Lagoze, Dean Krafft, Richard Marisa, John Saylor, Carol Terrizzi and Herbert Van de Sompel. 2002. "A Spectrum of Interoperability: The Site for Science Prototype for the NSDL." *D-Lib Magazine* 8, no. 1. doi:10.1045/january2002-arms

Bhandary, Priyanka, Arun S. Seetharam, Zebulun W. Arandsee, Manhoi Hur and Eve Syrkin Wurtle. 2018. "Raising Orphans from a Metadata Morass: A Researcher's Guide to Re-Use of Public 'omics Data." *Plant Science* 267: 32-47. doi:10.1016/j.plantsci.2017.10.014

Bean, Carol A. and Rebecca Green eds. 2001. *Relationships in the Organization of Knowledge*. Boston, MA: Kluwer.

Bevington, David, ed. 2019. *Hamlet: By William Shakespeare*. Internet Shakespeare Editions. https://internetshakespeare.uvic.ca/Library/Texts/Ham/

Boellstorff, Tom. 2013. "Making Big Data, in Theory." *First Monday* 18, no. 10. http://journals.uic.edu/ojs/index.php/fm/article/view/4869/3750

Borgman, Christine L. 2003. "The Invisible Library: Paradox of the Global Information Infrastructure." *Library Trends* 51, no. 4: 652-74.

Borgman, Christine L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT Press.

Borgman, Christine L., Jillian C. Wallis and Matthew S. Mayernik. 2012. "Who's Got the Data? Interdependencies in Science and Technology Collaborations." *Computer Supported Cooperative Work* 21, no. 6: 485-523. doi:10.1007/s10606-012-9169-z

Bowker, Geoffrey. 2016. "Just What are we Archiving?" *Limn* 6. http://limn.it/just-what-are-we-archiving/

Brodeur, Jean, Serena Coetzee, David Danko, Stephane Garcia and Jan Hjelmager. 2019. "Geographic Information Metadata: An Outlook from the International Standardization Perspective." *ISPRS International Journal of Geo-Information* 8: 280. doi:10.3390/ijgi8060280

Brunton, Finn. 2016. "Keeping the Books." *Limn* 6. https://limn.it/articles/keeping-the-books/

Buckland, Michael K. 1997. "What is a 'Document'?" *Journal of the American Society for Information Science* 48, no. 9: 804-9. doi:10.1002/(SICI)1097-4571(199709) 48:9<804::AID-ASI5>3.0.CO;2-V

Buckland, Michael. 2014. "Documentality Beyond Documents." *Monist* 97, no. 2: 179-86. doi:10.5840/monist 201497212

Buckland, Michael. 2017. *Information and Society*. Cambridge, MA: MIT Press.

Buckland, Michael. 2018. "Document Theory." *Knowledge Organization* 45: 425-36.

Cairns, Susan. 2013. "Mutualizing Museum Knowledge: Folksonomies and the Changing Shape of Expertise." *Curator: The Museum Journal* 56, no. 1: 107-19. doi:10.1111/cura.12011

Caplan, Priscilla. 2003. *Metadata Fundamentals for All Librarians*. Chicago: American Library Association.

Chamberlain, Alan and Andy Crabtree. 2016. "Searching for Music: Understanding the Discovery, Acquisition, Processing and Organization of Music in a Domestic Setting for Design." *Personal and Ubiquitous Computing* 20, no. 4: 559-71. doi:10.1007/s00779-016-0911-2

Coyle, Karen. 2010. "Library Data in a Modern Context." *Library Technology Reports* 46, no. 1: 5-13.

Crawford, Kate. 2019. "Halt the Use of Facial-Recognition Technology Until it is Regulated." *Nature* 572, no. 7771: 565. doi:10.1038/d41586-019-02514-7

Danko, David M. 2012. "Geospatial Metadata." In *Springer Handbook of Geographic Information,* ed. Wolfgang Kresse and David M. Danko. Berlin: Springer: 359-91. doi:10.1007/978-3-540-72680-7_12

Day, Ronald E. 2019. *Documentarity: Evidence, Ontology, and Inscription*. Cambridge, MA: MIT Press.

Dervin, Brenda. 1997. "Given a Context by Any Other Name: Methodological Tools for Taming the Unruly Beast." In *Information Seeking in Context,* ed. P. Vakkari, R. Savolainen and B. Dervin. London: Taylor Graham: 13-38.

Dourish, Paul. 2004. "What We Talk About When We Talk About Context." *Personal and Ubiquitous Computing,* 8, no. 1: 19-30. doi:10.1007/s00779-003-0253-8

Dunsire, Gordon, Dianne Hillmann and Jon Phipps. 2012. "Reconsidering Universal Bibliographic Control in light of the Semantic Web." *Journal of Library Metadata*, 12, no. 2-3: 164-76. doi:10.1080/19386389.2012.699831

Edwards, Paul N. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: MIT Press.

Edwards, Paul N. 2017. "Knowledge Infrastructures for the Anthropocene." *The Anthropocene Review* 4, no. 1: 34-43. doi:10.1177/2053019616679854

Edwards, Paul N., Matthew S. Mayernik, Archer L. Batcheller, Geoffrey C. Bowker and Christine L. Borgman. 2011. "Science Friction: Data, Metadata, and Collaboration." *Social Studies of Science* 41, no. 5: 667-90. doi:10.1177/0306312711413314

Fegraus, Eric H., Sandy Andelman, Matthew B. Jones and Mark Schildhauer. 2005. "Maximizing the Value of Ecological Data With Structured Metadata: An Introduction to Ecological Metadata Language (EML) and Principles for Metadata Creation." *Bulletin of the Ecological Society of America* 86, no. 3: 158-68. doi:10.1890/0012-9623(2005)86[158:MTVOED]2.0.CO;2

Feinberg, Melanie. 2017. "The Value of Discernment: Making Use of Interpretive Flexibility in Metadata Generation and Aggregation." *Information Research* 22, no. 1, CoLIS paper 1649. http://informationr.net/ir/22-1/colis/colis1649.html

Feinberg, Melanie. 2018. "Factotem: What is Information Access for?" *Cataloging & Classification Quarterly* 56, no. 8: 665-82. doi:10.1080/01639374.2018.1494654

Fidler, Bradley and Amelia Acker. 2016. "Metadata, Infrastructure, and Computer-Mediated Communication in Historical Perspective." *Journal of the Association for Information Science and Technology* 68, no. 2: 412-22. doi:10.1002/asi.23660

Foulonneau, Muriel and Jenn Riley. 2008. *Metadata for Digital Resources: Implementation, Systems Design and Interoperability*. Oxford: Chandos.

Furner, Jonathan. 2016. "'Data': The Data." In *Information Cultures in the Digital Age: A Festschrift in honor of Rafael Capurro*, ed. M. Kelly and J. Bielby. Wiesbaden: Springer, 287-306. doi:10.1007/978-3-658-14681-8_17

Furner, Jonathan. 2017. "Philosophy of Data: Why?" *Education for Information* 33, no. 1: 55-70. doi:10.3233/EFI-170986

Furner, Jonathan. 2020. "Definitions of 'Metadata': A Brief Survey of International Standards." *Journal of the Association for Information Science and Technology* 71: E-33-42. doi:10.1002/asi.24295

Gartner, Richard. 2016. *Metadata: Shaping Knowledge from Antiquity to the Semantic Web*. Cham: Springer. doi:10.1007/978-3-319-40893-4

Gazan, Rich. 2003. "Metadata as a Realm of Translation: Merging Knowledge Domains in the Design of an Environmental Information System." *Knowledge Organization* 30: 182-90.

Gil, Yolanda, Daniel Garijo, Varun Ratnakar, Deborah Khider, Julien Emile-Geay and Nicholas McKay. 2017. "A Controlled Crowdsourcing Approach for Practical Ontology Extensions and Metadata Annotations." In *The Semantic Web – ISWC 2017: 16th International Semantic Web Conference, Vienna, Austria, October 21-25, 2017, Proceedings*, ed. Claudia d'Amatop, Miriam Fernandez, Valentina Tamma, Freddy Lecue, Philippe Cudré-Mauroux, Juan Sequeda, Christoph Lange and Jeff Heflin. Lecture Notes in Computer Science 10588. Cham: Springer, 2: 231-46. doi:10.1007/978-3-319-68204-4_24

710

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

Giles, Jeremy R.A. 2011. "Geoscience Metadata—No Pain, No Gain." *Geological Society of America Special Papers*, vol. 482. Boulder, CO: Geological Society of America, 29-33. doi:10.1130/2011.2482(03)

Gilliland, Anne J. 2008. "Setting the Stage." In *Introduction to Metadata,* ed. Murtha Baca. 1st ed. Los Angeles, CA: Getty Information Institute. https://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.pdf

Gilliland, Anne J. 2014. *Conceptualizing 21st-Century Archives*. Chicago: Society of American Archivists.

Gitelman, Lisa ed. 2013. *"Raw Data" is an Oxymoron.* Cambridge, MA: MIT Press.

Gnoli, Claudio. 2008. "Ten Long-Term Research Questions in Knowledge Organization." *Knowledge Organization* 35: 137-49. doi:10.5771/0943-7444-2008-2-3-137

Gnoli, Claudio. 2012. "Metadata About What? Distinguishing Between Ontic, Epistemic, and Documental Dimensions in Knowledge Organization." *Knowledge Organization* 39: 268-75. doi:10.5771/0943-7444-2012-4-268

Gordon, Sean and Ted Habermann. 2018. "The Influence of Community Recommendations on Metadata Completeness." *Ecological Informatics* 43: 38-51. doi:10.1016/j.ecoinf.2017.09.005

Gorman, Michael. 1999. "Metadata or Cataloguing? A False Choice." *Journal of Internet Cataloging* 2, no. 1: 5-22. doi:10.1300/J141v02n01_03

Gorman, Michael. 2006. "Metadata Dreaming." *The Serials Librarian* 51, no. 2: 47-54.

Gorman, Michael. 2011. *Broken Pieces: A Library Life, 1941-1978*. Chicago: American Library Association.

Green, Rebecca. 2008. "Relationships in Knowledge Organization." *Knowledge Organization* 35: 150-9. doi:10.5771/0943-7444-2008-2-3-150

Greenberg, Jane. 2001. "A Quantitative Categorical Analysis of Metadata Elements in Image-Applicable Metadata Schemas." *Journal of the American Society for Information Science and Technology* 52, no. 11: 917-24. doi:10.1002/asi.1170

Greenberg, Jane. 2003. "Metadata and the World Wide Web." In *Encyclopedia of Library and Information Science*. 2nd ed., ed. Miriam A. Drake. New York: Marcel Dekker, 1876-88.

Greenberg, Jane. 2004. "Metadata Extraction and Harvesting: A Comparison of Two Automatic Metadata Generation Applications." *Journal of Internet Cataloging* 6, no. 4: 59-82. doi:10.1300/J141v06n04_05

Greenberg, Jane. 2005. "Understanding Metadata and Metadata Schemes." *Cataloging & Classification Quarterly* 40, no. 3: 17-36. doi:10.1300/J104v40n03_02

Greenberg, Jane. 2009. "Metadata and Digital Information." In *Encyclopedia of Library and Information Sciences*. 3rd ed., ed. Marcia J. Bates and Mary Niles Maack. Boca Raton, FL: Taylor & Francis, 3610-23.

Greenberg, Jane. 2017. "Big Metadata, Smart Metadata, and Metadata Capital: Toward Greater Synergy Between Data Science and Metadata." *Journal of Data and Information Science* 2, no. 3: 19-36. doi:10.1515/jdis-2017-0012

Greenberg, Jane and Emmanouel Garoufallou. 2013. "Change and a Future for Metadata." In *Metadata and Semantics Research. MTSR 2013,* ed. Emmanouel Garoufallou and Jane Greenberg, Communications in Computer and Information Science 390. Cham, Springer, 1-5. doi:10.1007/978-3-319-03437-9_1

Greenberg, Jane, Angela Murillo, Adrian Ogletree, Rebecca Boyles, Negin Martin and Charles Romeo. 2014. "Metadata Capital: Automating Metadata Workflows in the NIEHS Viral Vector Core Laboratory." In *Metadata and Semantics Research*, ed. Sissi Closs, Rudi Studer, Emmanouel Garoufallou, Miguel-Angel Sicilia 478. Cham, Springer, 1-13. doi:10.1007/978-3-319-13674-5_1

Habermann, Ted. 2018. "Metadata Life Cycles, Use Cases and Hierarchies." *Geosciences* 8, no. 5: 179. doi:10.3390/geosciences8050179

Han, Myung-Ja and Patricia Hswe. 2010. "The Evolving Role of the Metadata Librarian." *Library Resources & Technical Services* 54, no. 3: 129-41. doi:10.5860/lrts.54n3.129

Harkness Connell, T. and R.L. Maxwell eds. 2000. *The Future of Cataloging: Insights from the Lubetzky Symposium*. Chicago: American Library Association.

Haynes, David. 2017. *Metadata for Information Management and Retrieval*. 2nd ed. London: Facet Publishing.

Heritage, John. 1984. *Garfinkel and Ethnomethodology*. Cambridge, MA: Polity Press.

Hjørland, Birger. 2008. "What is Knowledge Organization (KO)?" *Knowledge Organization* 35: 86-101. doi:10.5771/0943-7444-2008-2-3-86

Hjørland, Birger. 2018. "Data (With Big Data and Database Semantics)." *Knowledge Organization* 45: 685-708.

Howarth, Lynne C. 2005. "Metadata and Bibliographic Control: Soul-Mates or Two Solitudes?" *Cataloging & Classification Quarterly* 40, no. 3-4: 37-56. doi:10.1300/J104v40n03_03

International Conference on Cataloguing Principles. 1971. *Statement of Principles: Adopted at the International Conference on Cataloguing Principles, Paris, October 1961*, ed. E. Verona. London: British Museum; International Federation of Library Associations (Committee on Cataloguing).

JSC (Joint Steering Committee for Development of RDA). 2014. *Resource Description & Access: RDA*. 2014 Revision. Chicago: American Library Association.

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

711

Jones, Matthew B., Schildhauer, Mark P., Reichman, O.J. and Bowers, Shawn. 2006. "The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere." *Annual Review of Ecology, Evolution, and Systematics* 37: 519-44. doi:10.1146/annurev.ecolsys.37.091305.110031

Kaase, Max. 2001. "Databases, Core: Political Science and Political Behavior". In *International Encyclopedia of the Social and Behavioral Sciences*, eds. Neil J. Smelser and Paul B. Baltes. Amsterdam: Elsevier, 3251–3255.

Khoo, Michael and Catherine Hall. 2013. "Managing Metadata: Networks of Practice, Technological Frames, and Metadata Work in a Digital Library." *Information and Organization* 23, no. 2: 81-106. doi:10.1016/j.infoandorg.2013.01.003

Klensin, John C. 1995. "When the Metadata Exceed the Data: Data Management with Uncertain Data." *Statistics and Computing* 5, no. 1: 73-84. doi:10.1007/BF00140667

Lagoze, Carl. 1996. "The Warwick Framework: A Container Architecture for Aggregating Sets of Metadata." *D-Lib Magazine*, 2, no. 7. http://www.dlib.org/dlib/july96/lagoze/07lagoze.html

Lagoze, Carl J. 2010. "Lost Identity: The Assimilation of Digital Libraries into the Web." Ph.D. diss., Cornell University.

Lagoze, Carl, Dean Krafft, Tim Cornwell, Naomi Dushay, Dean Eckstrom and John Saylor. 2006. "Metadata Aggregation and 'Automated Digital Libraries': A Retrospective on the NSDL Experience." In *Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries – JCDL '06.* New York: ACM Press, 230-239. doi:10.1145/1141753.1141804

Lahti, Leo, Jani Marjanen, Hege Roivainen and Mikko Tolonen. 2019. "Bibliographic Data Science and the History of the Book (c. 1500-1800)." *Cataloging & Classification Quarterly* 57, no. 1: 5-23. doi:10.1080/01639374.2018.1543747

Law, John. 2004. *After Method: Mess in Social Science Research.* New York: Routledge.

Lawrence, B.N., R. Lowry, P. Miller, H. Snaith and A. Woolf. 2009. "Information in Environmental Data Grids." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 367, no. 1890: 1003-14. doi:10.1098/rsta.2008.0237

Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study.* Chicago: University of Chicago Press.

Madrigal, Alexis C. 2014. "How Netflix Reverse Engineered Hollywood." *Atlantic*, Jan. 2.

Marchionini, Gary. 2006. "Exploratory Search: From Finding to Understanding." *Communications of the ACM* 49, no. 4: 41-6. doi:10.1145/1121949.1121979

Marchionini, Gary. 2012. "Award of Merit Acceptance Speech: Bridges, Linchpins and Membranes: From I to We." *Bulletin of the American Society for Information Science and Technology* 38, no. 2: 19-21. doi:10.1002/bult.2012.1720380206

Maron, Deborah and Erin Carter. 2017. "'More Than What It Seems': How Critical Theory, Popular Engagement and Apps Like Tinder Can Help Us Reframe Metadata and its Consequences." In *Advancing Metadata Practice: Quality, Openness, Interoperability 2017 Proceedings of the International Conference on Dublin Core and Metadata Applications 26-29 October 2017.* Silver Spring, MD: Dublin Core Metadata Initiative, 1-12. https://dcpapers.dublincore.org/pubs/article/view/3849

Mayer-Schonberger, Viktor and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think.* Boston: Houghton Mifflin Harcourt.

Mayernik, Matthew S. 2016. "Research Data and Metadata Curation as Institutional Issues." *Journal of the Association for Information Science and Technology* 67, no. 4: 973-93. doi:10.1002/asi.23425

Mayernik, Matthew S. 2018. "Scholarly Resource Linking: Building Out a 'Relationship Life Cycle.'" In *Proceedings of the 81st Annual Meeting of the Association for Information Science and Technology (ASIS&T)*, ed. Luanne Freund. Somerset, NJ: Wiley, 337-46. doi:10.1002/pra2.2018.14505501037

Mayernik, Matthew S. 2019. "Metadata Accounts: Achieving Data and Evidence in Scientific Research." *Social Studies of Science* 49, no. 5: 732-57. doi:10.1177/0306312719863494

Mayernik, Matthew S. and Amelia Acker. 2018. "Tracing the Traces: The Critical Role of Metadata Within Networked Communications." *Journal of the Association for Information Science and Technology* 69, no. 1: 177-80. doi:10.1002/asi.23927

Michener, William K., James W. Brunt, John J. Helly, T.B. Kirchner and Susan G. Stafford. 1997. "Nongeospatial Metadata for the Ecological Sciences." *Ecological Applications* 7, no. 1: 330-42. doi:10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2

Montenegro, María. 2019. "Subverting the Universality of Metadata Standards: The TK Labels as a Tool to Promote Indigenous Data Sovereignty." *Journal of Documentation*, 75, no. 4: 731-49. doi:10.1108/JD-08-2018-0124

Moulaison Sandy, Heather and Felicity Dykas. 2016. "High-Quality Metadata and Repository Staffing: Perceptions of United States-Based OpenDOAR Participants." *Cataloging & Classification Quarterly* 54, no. 2: 101-16. doi:10.1080/01639374.2015.1116480

712

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

Mühling, Markus, Manja Meister, Nikolaus Korfhage, Jörg Wehling, Angelika Hörth, Ralph Ewerth and Bernd Freisleben. 2019. "Content-Based Video Retrieval in Historical Collections of the German Broadcasting Archive." *International Journal on Digital Libraries* 20, no. 2: 167-83. doi:10.1007/s00799-018-0236-z

Padilla, Thomas. 2019. *Responsible Operations: Data Science, Machine Learning, and AI in Libraries*. Dublin, OH: OCLC Research. doi:10.25333/xk7z-9g97

Palmer, Carole L., Cheryl A. Thompson, Karen S. Baker and Megan Senseney. 2014. "Meeting Data Workforce Needs: Indicators Based on Recent Data Curation Placements." In *iConference 2014 Papers*. Champaign-Urbana, IL: University of Illinois, 522-37. doi:10.9776/14133

Park, Jung-ran and Susan Maszaros. 2009. "Metadata Object Description Schema (MODS) in Digital Repositories: An Exploratory Study of Metadata Use and Quality." *Knowledge Organization* 36: 46-59. doi:10.5771/0943-7444-2009-1-46

Pomerantz, Jeffrey. 2015. *Metadata*. Cambridge, MA: MIT Press.

Rasmussen, Karsten Boye. 2014. "Social Science Metadata and the Foundations of the DDI." *IASSIST Quarterly* 37, no. 1-4: 28-35. doi:10.29173/iq499

Renear, Allen H. and Karen M. Wickett. 2010. "There are No Documents." In *Proceedings of Balisage: The Markup Conference 2010*. Balisage Series on Markup Technologies. Rockville, MD: Mulberry Technologies. doi:10.4242/BalisageVol5.Renear01

Riley, Jenn. 2017. *Understanding Metadata: What is Metadata, and What is it For? A Primer*. Baltimore: National Information Standards Organization (NISO). https://www.niso.org/publications/understanding-metadata-2017

Rosenberg, Daniel. 2013. "Data Before the Fact." In *"Raw Data" Is an Oxymoron*, ed. Lisa Gitelman. Cambridge, MA: MIT Press, 15-40.

Schneider, David P., Clara Deser, John Fasullo and Kevin E. Trenberth. 2013. "Climate Data Guide Spurs Discovery and Understanding." *Eos, Transactions American Geophysical Union* 94, no. 13: 121-2. doi:10.1002/2013EO130001

Schneier, Bruce. 2014. "Metadata = Surveillance." *IEEE Security & Privacy* 12, no. 2: 84. doi:10.1109/MSP.2014.28

Schneier, Bruce. 2015. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York: Norton.

Schottlaender, Brian E.C. ed. 1998. *The Future of the Descriptive Cataloging Rules: Papers from the ALCTS Preconference, AACR2000, American Library Association Annual Conference, Chicago, June 22, 1995*. Chicago: American Library Association.

Seeman, Dean. 2012. "Naming Names: The Ethics of Identification in Digital Library Metadata." *Knowledge Organization* 39: 325-31. doi:10.5771/0943-7444-2012-5-325

Sicilia, Miguel-Angel ed. 2014. *Handbook of Metadata, Semantics and Ontologies*. Singapore: World Scientific Publishing.

Sinclair, James and Michael Cardew-Hall. 2007. "The Folksonomy Tag Cloud: When is it Useful?" *Journal of Information Science* 34, no. 1: 15-29. doi:10.1177/0165551506078083

Sisario, Ben. 2019. "In Streaming Age, Classical Music Gets Lost in the Metadata." *New York Times*, June 23.

Smiraglia, Richard P. 2005. "Introducing Metadata." In *Metadata: A Cataloger's Primer*, ed. R.P. Smiraglia. New York: Routledge, 1-15.

Srinivasan, Ramesh, Robin Boast, Jonathan Furner and Katherine Becvar. 2009. "Digital Museums and Diverse Cultural Knowledges: Moving Past the Traditional Catalog." *The Information Society* 25, no. 4: 265-78. doi:10.1080/01972240903028714

Srinivasan, Ramesh and Jeffrey Huang. 2005. "Fluid Ontologies for Digital Museums." *International Journal on Digital Libraries*, 5, no. 3: 193-204. doi:10.1007/s00799-004-0105-9

Sugimoto, Shigeo, Thomas Baker and Stuart L. Weibel. 2002. "Dublin Core: Process and Principles." In *Digital Libraries: People, Knowledge, and Technology: 5th International Conference on Asian Digital Libraries, ICADL 2002 Singapore, December 11–14, 2002 Proceedings*, ed. Ee- Peng Lim, Schubert Foo, Chris Khoo, Hsinchun Chen, Edward Fox and Shalini Urs. Lecture Notes in Computer Science 2555. Berlin: Springer, 25-35. doi:10.1007/3-540-36227-4_3

Svenonius, Elaine. 2000. *The Intellectual Foundation of Information Organization*. Cambridge, MA: MIT Press.

Szostak, Rick. 2012. "Toward a Classification of Relationships." *Knowledge Organization* 39: 83-94. doi:10.5771/0943-7444-2012-2-83

Talja, Sanna, Heidi Keso and Tarja Pietiläinen. 1999. "The Production of 'Context' in Information Seeking Research: A Metatheoretical View." *Information Processing & Management* 35, no. 6: 751-63. doi:10.1016/S0306-4573(99)00024-2

Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff and Mike Frame. 2011. "Data Sharing by Scientists: Practices and Perceptions." *PLoS ONE* 6, no. 6: e21101. doi:10.1371/journal.pone.0021101

Tillett, Barbara B. 2003. "AACR2 and Metadata: Library Opportunities in the Global Semantic Web." *Cataloging & Classification Quarterly* 36, no. 3-4: 101-19. doi:10.1300/J104v36n03_09

Knowl. Org. 47(2020)No.8
M. S. Mayernik. Metadata

713

Urban, Richard J. 2014. "The 1:1 Principle in the Age of Linked Data." In *Metadata Intersections: Bridging the Archipelago of Cultural Memory: 2014 Proceedings of the International Conference on Dublin Core and Metadata Applications Proceedings,* ed. William Moen and Amy Rushing. Silver Spring, MD: Dublin Core Metadata Initiative, 119-28. https://dcpapers.dublincore.org/pubs/article/view/3707/0

Van de Sompel, Herbert, Michael L. Nelson, Carl Lagoze and Simeon Warner. 2004. "Resource Harvesting Within the OAI-PMH Framework." *D-Lib Magazine* 10, no. 12. doi:10.1045/december2004-vandesompel

Vertesi, Janet, Jofish Kaye, Samantha N. Jarosewski, Vera D. Khovanskaya and Jenna Song. 2016. "Data Narratives: Uncovering Tensions in Personal Data Management." In *CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing.* New York: ACM Press, 478-90. http://doi.org/10.1145/2818048.2820017

Weibel, Stuart. 1995. "Metadata: The Foundations of Resource Description." *D-Lib Magazine* 1, no. 1. http://www.dlib.org/dlib/July95/07weibel.html

Weihs, Jean ed. 1998. *The Principles and Future of AACR: Proceedings of the International Conference on the Principles and Future Development of AACR.* Chicago: American Library Association.

White, Hollie C. 2010. "Considering Personal Organization: Metadata Practices of Scientists." *Journal of Library Metadata* 10, no. 2: 156-72. doi:10.1080/19386389.2010.506396

Wickett, Karen M. 2015. "Accounting for Context in Markup: Which Situation, Whose Semantics?" In *Proceedings of Balisage: The Markup Conference 2015.* Balisage Series on Markup Technologies Rockville, MD: Mulberry Technologies. doi:10.4242/BalisageVol15.Wickett01

Wickett, Karen. 2018. "A Logic-Based Framework for Collection/Item Metadata Relationships." *Journal of Documentation* 74, no. 6: 1175-89. doi:10.1108/JD-01-2018-0017

Woolgar, Steve. 1981. "Critique and Criticism: Two Readings of Ethnomethodology." *Social Studies of Science* 11, no. 4: 504-14. doi:10.1177/030631278101100406

Zeng, Marcia Lei. 2019. "Interoperability." *Knowledge Organization* 46: 122-46.

Zeng, Marcia Lei and Jian Qin. 2016. *Metadata.* 2nd ed. Chicago: Neal-Schuman.