Found in Translation: Four Characteristics of Metadata Practice

Martin Kurth

Introduction[1]

In the growing body of metadata literature, the practitioners, researchers, and educators contributing to it have addressed metadata practice almost exclusively in terms of the metadata itself. This attention to the objects of metadata work has certainly been necessary at such an early stage in the development of metadata practice, in order to allow those engaged in it to share and establish optimum strategies for creating and manipulating metadata that best serves the users of the resources it describes. But perhaps now is an opportune time to step back from a metadata-centric approach in order to shift focus away from the objects of metadata work and instead toward metadata practice as practice, as a specialty that has distinct characteristics. Exploring the work of metadata practitioners in this way and at this time is warranted for three primary reasons. First, as already noted, the metadata literature is maturing and thus offers a rich resource upon which to build a study of metadata practice; second, digital resources and the metadata used to access and manage them are continuing to grow in importance to information seekers worldwide; and, third, because of this growth, more

[1] It is a privilege to have the opportunity to discuss metadata practice in a festschrift to honor Tom Turner. In addition to being an outstanding metadata librarian, Tom was a great champion of metadata work because he loved to think about it and talk about it to anyone interested in learning more. My only regret in writing this essay is that I was not able to discuss it with Tom because it would have undoubtedly been better as a result.

and more workers in libraries and other information service organizations are engaging in metadata practice.[2]

A discussion of metadata practice presents an opportunity to reflect on some of its significant characteristics: on the approaches that practitioners typically use when they work with metadata, on the layered responsibilities of metadata practitioners in the collaborative environment in which they work, on the fundamental role that metadata practitioners perform in all aspects of their work, and on the central contributions of metadata specialists toward interdisciplinary communication and collaboration. The primary purpose of presenting these significant features of metadata practice is to begin a conversation about metadata practice that will identify metadata practitioners' unique contributions toward facilitating the use of digital information.

Discussing important aspects of metadata practice also serves a secondary aim. The reflections on metadata practice that follow draw on both my experience in the Metadata Services Unit of the Cornell University Library (CUL) and my work as a library cataloger and cataloging manager. My transition from cataloging to metadata work has made it possible for me to approach metadata practice in light of my cataloging experience. Because many catalogers have also made or will make the transition to working with non-catalog metadata, a discussion of metadata practice can help those catalogers relate it to cataloging practice. To that end, I discuss metadata practice in light of cataloging work in those instances where I see clear similarities or differences.

Given the aims just described, I hope this essay will engage three primary audiences. First and foremost, metadata practitioners may be able to use it to reflect on

[2] For a compendium that demonstrates the maturation of metadata literature, see: Diane I. Hillmann and Elaine L. Westbrooks, eds., *Metadata in Practice* (Chicago: American Library Association, 2004).

the skills they possess and to identify the ways in which they can best use those skills in multidisciplinary collaborations; second, catalogers may find it helpful in understanding how their cataloging skills apply to metadata work; and, third, library and information science educators and training providers may find it instructive as they develop and implement models of metadata practice that are informed by experience in a library metadata services unit. Because the experience of working with metadata in libraries shapes this essay to such a great extent, the brief section that follows offers a history of the CUL Metadata Services Unit as a basis for understanding the nature of that experience.

The Cornell Metadata Services Unit

The managers of CUL's Central Technical Services (CTS) department formed its Metadata Services Unit[3] in January 2002 as a step toward achieving the goal they set in a February 2001 planning workshop to "[p]osition CTS as a key player in CUL digital initiatives."[4] To achieve their desire that CTS assume a significant role in the library's digital future, CTS managers created a metadata unit in direct response to Objective II.3 of *Cornell University Library Goals and Objectives 2002-2007*, which was to "[e]stablish and operate a 'consulting to production' metadata service capable of producing metadata in a variety of formats to organize, manage, and preserve

[3] Cornell University Library, Metadata Services, http://metadata.library.cornell.edu/ (8 Mar. 2008).
[4] Karen Calhoun, "Results of CTS Future Search Workshop: A Planning Exercise (February 23, 2001)," http://www.library.cornell.edu/cts/futuresearch/futworkshop.html (8 Mar. 2008). For a fuller discussion of the planning process that drove the creation of the Metadata Services Unit, see: Karen Calhoun, "Technology, Productivity and Change in Library Technical Services," *Library Collections, Acquisitions, & Technical Services* 27 (2003): 285-87.

collections over time and to enable effective discovery and use."[5] At the time of this writing, the Metadata Services Unit consists of four and one half full-time-equivalent staff (including the unit head) with backgrounds in cataloging, acquisitions, and information technology. Though Metadata Services is organizationally part of CTS, it has also from its inception served as the primary metadata service provider in CUL's Digital Consulting and Production Services (DCAPS), which offers a modular suite of cost-recovery services to create and manage digital resources for clients inside and outside of Cornell University.[6] In addition to metadata services, DCAPS provides services in digital media creation and reformatting, electronic publishing, copyright, and technology support.

To fulfill its role in achieving CTS and DCAPS service objectives, Metadata Services has adopted this mission: "Metadata Services provides metadata consulting, design, development, production, and conversion services to Cornell's faculty, staff, and community partners to increase the value of their digital resources."[7] On our web site we define "metadata" for our clients in this way: "Metadata organizes information about digital resources, including titles, authors, keywords, format, versions, and rights. It increases the value of digital resources by making them easier to access, use, share, and re-purpose."[8] Given this environmental context, the present essay considers metadata practice in terms of the metadata that practitioners design, develop, produce, and convert to other formats in order to manage digital resources and make them accessible

[5] Cornell University Library, "GOAL II: Provide Digital 'Life-Cycle' Production Services," *Cornell University Library Goals and Objectives 2002-2007*, http://www.library.cornell.edu/Admin/goals/goal2.html (8 Mar. 2008).
[6] Cornell University Library, "Digital Consulting & Production Services," updated 25 Apr. 2007. http://dcaps.library.cornell.edu/ (8 Mar. 2008).
[7] Cornell University Library, Metadata Services.
[8] Cornell University Library, Metadata Services.

to end users.

Metadata Practice Approaches Metadata in Aggregates

Metadata work as it is practiced in libraries typically deals with metadata in aggregates, that is, in large groups of records that practitioners manipulate en masse.[9] Most frequently, practitioners approach metadata aggregates in terms of projects, collections, or, more broadly, metadata as it is used within specific communities of practice. Supporting these claims, guides to best practices for developing collections of digital materials reflect an aggregate-oriented view. The NISO *Framework of Guidance for Building Good Digital Collections*, for example, is emphatic in its holistic view of what it calls the third stage of digital collection development, characterized by the desirability of making resources available beyond known user communities: "Objects, metadata, and collections must now be viewed not only within the context of the projects that created them but as building blocks that others can reuse, repackage, and build services upon."[10] It is worth noting that although the NISO *Framework*'s principles foster the inter-collection interoperability of digital resources, its title, structure, and four core entities of collections, objects, metadata, and projects focus on the collection as the primary organizing construct for digital materials. Similarly, The *NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials* echoes the NISO *Framework*'s view of broad digital

[9] The *American Heritage Dictionary of the English Language, 4th ed.* definition of the noun "aggregate" is particularly relevant here: "A total considered with reference to its constituent parts."
[10] NISO Framework Advisory Group, *A Framework of Guidance for Building Good Digital Collections*, 2nd ed. (Bethesda, MD: National Information Standards Organization, 2004), 1, http://www.niso.org/framework/framework2.pdf (28 Nov. 2004).

integration by seeking to promote local digital collection-building practice that considers "the overall vision and goals for a networked cultural heritage."[11] The developers of the *NINCH Guide* argue that by adhering to community-wide best practices and a vision of cross-collection integration, digital collection builders "can ensure the quality, consistency and reliability of a project's digital resources and make them compatible with resources from other projects and domains, building on the work of others."[12] To promote the achievement of these goals, the NINCH Guide authors include among their six core principles of best practice the goals of optimizing the interoperability of resources across digital collections and of enabling the broadest use of resources by multiple audiences.[13] The NISO *Framework*'s principles for good digital collections agree with the *NINCH Guide*'s vision that a digital collection should integrate with important digital initiatives that relate to it.[14]

Given this perspective, namely that digital projects focus on collection building and interoperability among collections, metadata work scenarios typically begin by considering the goals, functional requirements, and user needs of the digital collection project at hand. They then consider the research and instructional community or communities that developers intend for the collection to serve, other significant digital initiatives serving those communities, and mechanisms for inter-collection interoperability such as federated searching or metadata harvesting that already may be in place. It is only after considering such features of "the big picture" that metadata

[11] National Initiative for a Networked Cultural Heritage, *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials* (National Initiative for a Networked Cultural Heritage, 2003), 3, http://www.nyu.edu/its/humanities/ninchguide/ (8 Mar. 2008).
[12] *NINCH Guide*, 1.
[13] *NINCH Guide*, 4.
[14] NISO *Framework*, 10.

practitioners begin to consider the structure and contents of the metadata records that will properly serve the needs of the collection. Metadata practitioners begin the process of ensuring that records will meet collection needs by developing record exemplars for representative digital objects in the collection.

This collection-to-record approach, which is typical in metadata work, differs from the inverse approach that is predominant in most library cataloging units. Catalogers are able to use a record-to-collection approach for the majority of library materials that cross their desks because the infrastructure for creating and delivering catalog records is much more established than the digital world's infrastructure, which has no content and encoding standards as pervasive as the International Standard Bibliographic Descriptions (ISBDs) and MARC 21 formats. In other words, a cataloger typically starts with a library resource to be cataloged, whether it is an electronic resource, book, serial, or DVD. Notwithstanding some exceptions that I will discuss below, the cataloger as a rule first considers the item at hand and describes it in a catalog record. From this record-centric perspective, the cataloger only then looks to connect the item cataloged to related items via classification numbers, subject headings, and authoritative name and title entries. That catalogers do not have to consider the catalog's context and structures before they catalog an item is thus one of the legacies of the well-documented theory and practice of the cataloging community. The digital resource metadata community, which in its best examples draws on the conceptual foundations and lessons learned from the cataloging community, does not enjoy the same heritage of delivery mechanisms (i.e., the various iterations of the library catalog), tools, and documentation that catalogers rely on in their typical record-to-collection workflows. Figure 1 seeks to represent graphically the differences between cataloging

and metadata approaches.

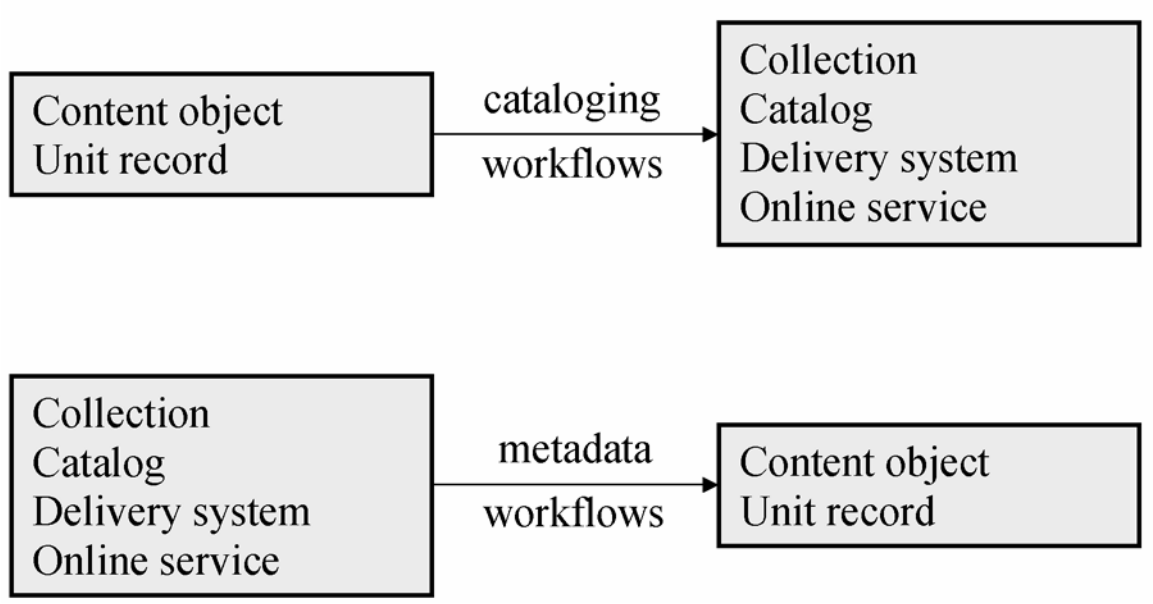

Figure 2.1 Cataloging vs. metadata workflows: An object focus vs. a collection focus

Though the above representation of cataloging and metadata workflows as inversions of one another is instructive, it needs some qualification. In manual cataloging workflows, exceptions to the rule do exist. For example, catalogers faced with large sets of items, such as titles in series, microform collections, archival collections, or classes of materials that the library has designated for minimal-level cataloging, typically consider the sets as a whole before they address individual objects within them. In the case of automated cataloging workflows, cataloging units accepting large sets of MARC records for microform collections or, more recently, for electronic journal aggregations, consider a record set as a whole and use automated scripts to make universal additions and deletions to records in the set. In fact, in a 2005 article, Cornell librarians David Banush, Jean Pajerek, and I characterized the history of electronic

journal cataloging as an evolution from manual record-by-record approaches to automated approaches that treat records for electronic journals in aggregated sets.[15] Seen in this light, the typical metadata collection-to-record workflow is similar to, or in some cases even an outgrowth of, the set-oriented cataloging workflows that automation and cooperative cataloging have increasingly made possible for cataloging units to use for large aggregations of library materials. Nevertheless, in spite of such convergence between set-oriented cataloging and metadata workflows, the predominant manual workflow in cataloging units reflects the record-to-collection model proposed here.

Why is it important to make this distinction between metadata and cataloging practice? Seeing metadata and cataloging workflows as inversions of one another can be useful for catalogers who are planning to become more involved in metadata work. It can help them see that their skills are still applicable in a metadata environment because the processes involved are related—they are in fact mirror images of one another.

That metadata practitioners usually approach metadata records in aggregates is thus one characteristic of metadata practice. To illustrate another important characteristic, the following section examines the collaborative nature of current metadata practice to identify interrelated layers of metadata responsibilities.

Metadata Practice Comprises Interpersonal, Informational, and Operational Layers

Metadata work to support digital collection creation and management rarely takes place in isolation; typically, metadata practitioners work collaboratively as

[15] David Banush, Martin Kurth, and Jean Pajerek, "Rehabilitating Killer Serials: An Automated Strategy for Maintaining E-Journal Metadata," *Library Resources and Technical Services Library Resources and Technical Services* 49:3 (2005): 190-203.

members of teams that bring together a variety of skill sets. The DCAPS model for digital production services mentioned above is representative of a collaborative approach that brings together digital media specialists, information technologists, metadata practitioners, and other library staff members with expertise relevant to digital collection development. In addition, digital collection teams often include subject domain specialists who have an intimate knowledge of the materials in the collection. In their study of digital library programs, Daniel Greenstein, then Director of the Digital Library Federation, and Suzanne Thorin, Dean of Indiana University Libraries, found such collaboration typical of digital collection efforts supporting higher education research and instruction; moreover, they argued that digital collection building is only possible "through the combined thinking of scholars, information professionals, and technologists."[16]

Typical of the collaborative nature of metadata practice, the mission statement of CUL Metadata Services identifies these activities that comprise metadata practice: consulting, design, development, production, and conversion. Expanding these five aspects of metadata practice into typical scenarios suggests the collaborative environment in which metadata practitioners work. Specifically, metadata staff consult with colleagues and clients regarding overall metadata requirements and constraints. Additionally, they work with subject specialists and user services librarians to design metadata record structures for digital collection access. They develop auxiliary resources such as application profiles, XML schemas, and XML namespaces to facilitate ongoing

[16] Daniel Greenstein and Suzanne E. Thorin, *The Digital Library: A Biography* (Washington, DC: Digital Library Federation, 2002), 31.

digital collection management by those responsible for it.[17] They produce metadata records to represent key features of digital resources. Finally, they convert metadata records from one scheme or format to another to support new uses or users.

This work, especially the consulting, design, and development components, is more akin to the work of managers than it is to the work of staff with no managerial responsibilities. In writing about information management in organizations, Chun Wei Choo, Professor of the Faculty of Information Studies at the University of Toronto, describes management as a series of conversations about commitments, especially creating, nurturing, or fulfilling them.[18] Much in the same way, metadata practitioners consulting with team members or clients engage in conversations that focus on commitments, either making, reporting on, modifying, or delivering the commitments that they and their metadata colleagues contribute to the digital collection effort. To carry the comparison of metadata practitioners to managers a step further, the digital project meeting room is also a rhetorical space in which metadata staff use their understanding of metadata requirements to negotiate with team members regarding specific courses of action.[19] In other words, such interpersonal interactions as negotiation and advocacy are common components of collaboration in metadata practice.

Because of the consequences that these aspects of collaboration have for

[17] For an introduction to application profiles, see: Rachel Heery and Manjula Patel, "Application Profiles: Mixing and Matching Metadata Schemas" *Ariadne* 25 (2000), http://www.ariadne.ac.uk/issue25/app-profiles/ (8 Mar. 2008). For a reasonably readable technical introduction to schemas and namespaces, see: Priscilla Walmsley, *Definitive XML Schema* (Upper Saddle River, NJ: Prentice Hall PTR, 2002), 3-8, 37-40.
[18] Chun Wei Choo, *Information Management for the Intelligent Organization*, 3rd ed. (Medford, NJ: Information Today), 61. In characterizing managerial practice as conversations about commitments, Choo makes use of: Terry Winograd and Fernando Flores, *Understanding Computers and Cognition: A New Foundation for Design* (Norwood, NJ: Ablex, 1986), 151.
[19] Choo, *Information Management*, 61. To establish the rhetorical nature of management, Choo draws on: Robert G. Eccles and Nitin Nohria, *Beyond the Hype: Rediscovering the Essence of Management* (Boston: Harvard Business School Press), 29-30.

metadata work, metadata practitioners perform multiple roles on any given digital collection project team. Again, it is useful to draw a parallel here to managerial practice. Choo, following McGill University management theorist Henry Mintzberg, identifies three roles that managers perform within their organizational groups: interpersonal roles, informational roles, and decisional roles.[20] According to Choo's categorization of these three types of management roles, managers' role of gathering and sharing information is pivotal in connecting their interpersonal relationships with colleagues to their commitment-related responsibilities in initiating projects and allocating resources. Similarly, with regard to metadata practice, a metadata practitioner's multiple roles break out into three facets that parallel a manager's roles, though as a practitioner the metadata specialist's commitment-related role is more operational than decisional. Figure 2 draws on the work of Choo and Mintzberg to illustrate the three-part layering of metadata practitioners' responsibilities to the organizations they serve. As with managers, metadata practitioners' informational roles tie their other roles together. For example, metadata specialists' research of relevant practices and standards informs their interpersonal expressions of metadata concerns to team members; in much the same way, their research shapes the operational choices they make as they develop and implement metadata element applications that address a digital collection's functional requirements.

[20] Choo, *Information Management*, 65-7. Henry Mintzberg, *The Nature of Managerial Work* (New York: Harper & Row, 1973), 54-94.

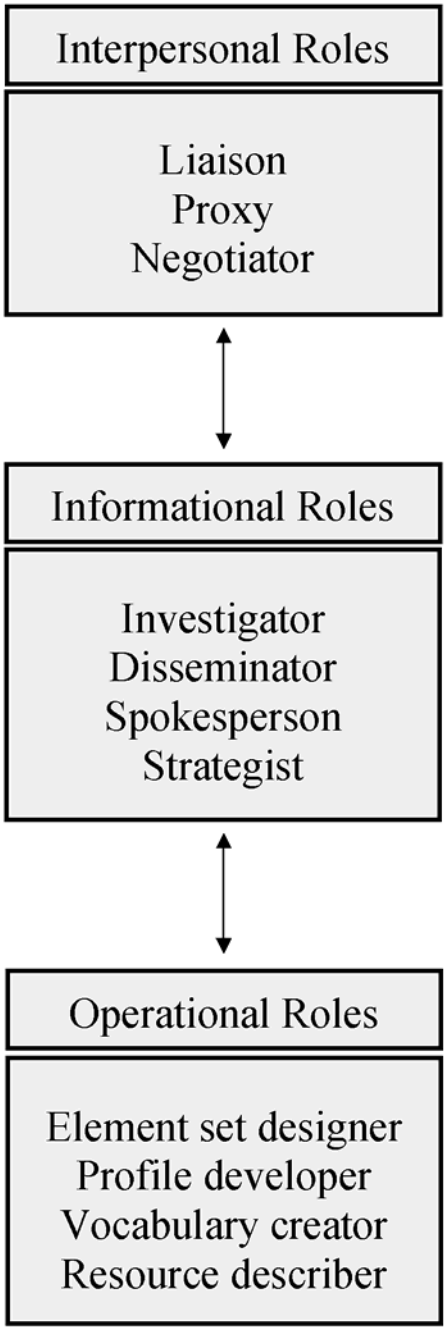

Figure 2.2. The metadata practitioner's organizational responsibilities
(Adapted from Mintzberg by way of Choo)[21]

If we accept that metadata practitioners' responsibilities in developing digital collections consist of interpersonal, informational, and operational layers or facets, then

[21] Choo, *Information Management*, 67. Mintzberg, *Nature of Managerial Work*, 59

we might reasonably ask whether the same could be said of all participants in digital collection projects. Do not digital media specialists, for example, have roles as liaisons and investigators, as well as operational roles as scanning and reformatting technicians? If metadata practitioners are not unique in having layers of complexity in their organizational responsibilities, what characteristic is unique to their work in designing, implementing, mapping, and converting metadata files, as well as to their interactions with their colleagues and clients concerning that work? Answering the latter question, which seeks to understand the primary feature of metadata practice, calls for a deeper analysis of the actions that underlie a metadata specialist's interpersonal, informational, and operational roles.

Metadata Practice Specializes in Cross-Community Translation

As previously noted, academic library digital collection efforts typically bring together specialists from different areas of the university. Libraries contribute the time of digital media specialists, information technologists, metadata practitioners, copyright experts, and public services or collection development librarians having subject expertise. Instructional departments and research centers contribute the work of faculty members or research staff who are creators or users of the resources in the digital collections. To understand metadata practitioners' roles in such collaborations more fully, it is necessary to learn more about how members of groups perform specialized tasks in contemporary society. Specialization makes it possible for any complex society to be efficient in its various undertakings by sustaining communities of practice around work-related specialties.

Social learning theorist Etienne Wenger has developed the term "community of

practice" in the following way: Human beings pursue enterprises in groups, learning together as they work together. From such group learning, they develop common practices that support both their shared work and the social interactions that facilitate it. As these practices coalesce over time, group members become assets that the group nurtures. Wenger calls groups who develop such shared practices "communities of practice."[22] Wenger argues that communities of practice serve four primary functions within complex societies. First, they enable members to create meaning by making available a history of practice that current members consult and interpret in order to build on it. Second, they preserve and create group knowledge that yields experienced, competent members who sustain the community's vitality into the future. Third, they spread information by making it a requirement for participation and by providing mechanisms for all members to share in new information obtained by an individual member. Finally, they create group-related identities for members by enabling them to invest their efforts in the group and by recognizing the abilities that they contribute to the group.[23]

Certainly metadata practitioners, and to a greater extent catalogers, can envision themselves as members of a community of practice supported by local policies and procedures; educational and training programs; standards, rules, codes, and reference tools; associations, conferences, and meetings; and the acronym-filled language by which we identify each other as members of the group. In the same way, specialists with whom we collaborate also have communities that influence and support their work. But what is important, if not exceptional, about metadata practitioners' roles as participants

[22] Etienne Wenger, *Communities of Practice: Learning, Meaning, and Identity* (Cambridge: Cambridge University Press, 1998), 45.
[23] Wenger, *Communities*, 251-52.

in collaborative enterprises that engage members from several communities of practice? Answering that question requires us to focus on the drawbacks, rather than the benefits, of specialization.

Though community-based specialization increases efficiency within societies in the ways identified by Wenger, it also creates barriers between specialties.[24] A community's shared culture and language facilitate communication within it, but at the same time make it more difficult for group members to discuss the group's activities outside the group. In this way, a group's conceptual framework and language erect communication boundaries around the perimeter of the group.[25] Because communities have developed their own languages for use within group boundaries, individuals responsible for communicating outside the group find it necessary to recode information for others to understand it. Communicating across boundaries requires learning the language and cultural framework of the group to which the information is to be passed. People who successfully communicate across group boundaries are thus able to negotiate successfully among two or more communities of practice.

Such boundary spanning is a two-step process: First, the person communicating across the community boundary gathers information from one side of the boundary and, second, he or she translates it to meet the cultural and language requirements of the other group. To investigate these boundary spanning behaviors, management researchers Tushman and Scanlan conducted a study of communication patterns in research and development groups in high-tech corporations. Their research indicates

[24] Choo, *Information Management*, 159.
[25] Michael L. Tushman and Thomas J. Scanlan, "Boundary Spanning Individuals: Their Role in Information Transfer and Their Antecedents," *The Academy of Management Journal* 24, no. 2 (June 1981): 290.

that people who are successful at spanning community boundaries actively develop formal and informal information sources both inside and outside their primary work communities.[26] Successful boundary spanners use their internal and external information sources to support their work of translating specific messages across communities.

From this characterization of a boundary-spanning individual who uses formal and informal sources from several communities to translate information among communities, a connection to the work of the metadata practitioner begins to emerge. Metadata specialists perform community-spanning translations at all levels of their work. At the interpersonal level, for example, they seek to reconcile the types of searches subject experts want to perform on digital collections with delivery system limitations identified by information technologists and with metadata requirements for interoperability with related collections. Metadata practitioners also perform inter-community translations at the informational level of their work when they investigate controlled vocabularies that relate to the terminology used in the digital collections they help to create. They carry these translations yet further at the operational level by developing element-set application profiles and dictionaries of locally used controlled vocabulary terms that address the needs of users within the constraints of the delivery platform. Ultimately, they may perform additional operational translations to map and transform metadata from local collections to make it available for harvesting and reuse in other online services.

Seen in this way, metadata practice conducts interrelated translations that relay

[26] Tushman and Scanlan, "Boundary Spanning Individuals," 292-93, 299-300.

messages between communities of practice in order to build collections and information systems that support the endeavors of those communities. In other words, the work of metadata practitioners takes place within the context of their own community of metadata practice; moreover, metadata practitioners use their interpersonal roles as liaisons, their informational roles as investigators, and their operational roles as element designers and vocabulary creators to gain access to the communities of practice of their digital project collaborators.[27] They do this by actively engaging with the language that those communities use. By working with members of other communities of practice to refine and apply their language in information systems, metadata practitioners contribute to the process of meaning-making that is crucial to community participation and that enables the metadata practitioners to become specialized participants in those communities.[28]

As such specialized participants, the responsibilities of metadata practitioners are twofold. First, they help the community to regularize its use of selected terminology to make it more accessible in information systems that may serve multiple communities. Second, they map community terminology to that of other communities to facilitate the exchange of information with those communities. Metadata practitioners bring their boundary-spanning translation skills to bear in fulfilling both of these responsibilities. In the former, they focus on intra-community communication to achieve inter- or cross-community ends; in the latter, they focus directly on cross-community communication.

[27] For a work that identifies metadata specialists as members of a community of practice, see: Norm Friesen, "Semantic Interoperability and Communities of Practice," 5 Feb. 2002, under "Language: Sign System or Social Process?" http://www.cancore.ca/documents/semantic.html (8 Mar. 2008).
[28] For the role of language in the process through which newcomers become participants in communities of practice, see: Jean Lave and Etienne Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge: Cambridge University Press, 1991), 105-9.

If we portray metadata practitioners as translators or interpreters who facilitate communication via information systems within and among communities of practice, does this perspective help us understand the day-to-day operational activities of metadata practitioners?[29] The following section explores the operational implications of viewing metadata practice as translation.

Metadata Practice's Semantic and Syntactic Translations Support Interoperability

Digital collection development efforts regularly incorporate a variety of metadata types designed to serve the interests of diverse groups.[30] For example, metadata for a digital object may contain descriptive metadata to serve the searching needs of end users. It may contain technical metadata to reflect the work and interests of digital imaging specialists. It may also contain preservation metadata that supports the goals of digital preservationists. And, it may contain rights metadata that addresses the concerns of intellectual property holders.

As encoding standards and content guidelines emerge for these various metadata types, metadata practitioners working on collection projects apply them to the resources at hand. At the same time, practitioners frequently inherit metadata from diverse sources in disparate formats, which they adapt and combine in light of operative community standards or guidelines.[31] Because a given metadata record may comprise components created by different groups at different times adhering to the parameters of

[29] For a discussion of "metadata creation as an interpretive undertaking," see: Friesen, "Semantic Interoperability," under "Example: Learning Resource Types" and "Language: Sign System or Social Process?"
[30] National Research Council, Committee on an Information Technology Strategy for the Library of Congress, *LC21: A Digital Strategy for the Library of Congress* (Washington, DC: National Academy Press, 2000), under "Metadata As a Cross-Community Activity," http://books.nap.edu/html/lc21/ (8 Mar. 2008).
[31] Godfrey Rust and Mark Bide, "The <indecs> Metadata Framework: Principles, Model, and Data Dictionary," June 2000. http://www.doi.org/topics/indecs/indecs_framework_2000.pdf (8 Mar. 2008), 6.

different communities, metadata records are typically modular aggregations of metadata expressions from multiple sources.[32]

The modularity that is widespread in metadata records is not foreign to library cataloging. A typical catalog record is similarly modular insofar as it may contain an International Standard Book Number assigned by a publisher, a Library of Congress Classification number assigned in one library, Library of Congress Subject Headings (LCSH) assigned in another library, and Anglo-American Cataloging Rules descriptive cataloging created in yet another. Nonetheless, the modularity of metadata used to support digital collections differs from catalog metadata in that its sources are more diverse, the standards for creating it are significantly less established and pervasive, and metadata practitioners approach it in aggregates.

Thus metadata practitioners typically reconcile legacy metadata with applicable standards, translate the language of one community to that of another, and balance the requirements of the collection at hand with the benefits of interoperating with related collections. It is within this richly nuanced—if not ambiguous—information environment that metadata specialists perform metadata practice's central operations of mapping and transformation.[33] Moreover, to extend the preceding sections' analyses, we can argue that mapping and transformation are the primary activities through which metadata practitioners engage in cross-community translation in the operational layer

[32] Rust and Bide, "The <indecs> Metadata Framework," 5. Diane Hillmann, Naomi Dushay, and Jon Phipps, "Improving Metadata Quality: Augmentation and Recombination," 2004, 4, http://ecommons.library.cornell.edu/bitstream/1813/7897/1/Paper_21.pdf (8 Mar. 2008). Heery and Patel, "Application Profiles," under "Background."

[33] For a discussion of the roles of mapping and transformation in metadata processing, see: Martin Kurth, David Ruddy, and Nathan Rupp, "Repurposing MARC Metadata: Using Digital Project Experience to Develop a Metadata Management Design," *Library Hi Tech* 22, no. 2 (2004): 153-59.

of their work.

To establish a baseline for the present discussion, it is useful to define mapping as "establishing relationships between semantically equivalent elements in different metadata schemes" and transformation as "the design and implementation of scripts and other tools that move mapped metadata between schemes."[34] Mapping typically involves creating tables, or crosswalks, that graphically represent the equivalencies among elements in different metadata schemes. Transformation typically requires automated processes that convert one metadata expression to another. For example, a metadata mapping and transformation scenario might translate a relational database of terms and identifiers used by classicists to describe digital images of ancient Greek inscriptions into Visual Resources Association (VRA) Core Categories metadata records encoded in XML.[35] The translation process might involve, first, mapping database fields containing the classicists' terms into VRA elements that contain terms used in the Art & Architecture Thesaurus® and, second, transforming the database rendering of those terms into the VRA structure called for in a VRA XML schema.

In fact, to use a cataloging analogy, this translation process of mapping and transforming parallels the activities that catalogers perform whenever they create catalog records for library resources. In their work, catalogers translate the language of the community that produced the resource to be cataloged into terms and syntax called for by the communities that sustain the ISBDs, MARC 21 formats, and any number of subject vocabularies.

[34] Kurth et al., "Repurposing MARC Metadata," 154, 157.
[35] Visual Resources Association Data Standards Committee, "VRA Core Categories, Version 3.0," 20 Feb. 2002. http://www.vraweb.org/resources/datastandards/vracore3/categories.html (8 Mar. 2008).

Authors writing about metadata mapping and transformation have often found it useful to discuss them in terms of semantics, the meaningful content that metadata conveys, and syntax, the structure that expresses that content.[36] Though mapping is indeed a semantic operation, it would be an oversimplification to view transformation as wholly syntactic. Rather, transformations involve translations that affect meaning as well as syntax because a target metadata structure may not be able to express the meaning of a source metadata expression exactly.[37] A cataloging example illustrates this point. The MARC syntax uses subfields to encode LCSH strings with a great deal of specificity. Though mapping might equate a MARC 650 topical subject term field with a Text Encoding Initiative (TEI) Guidelines for Electronic Text Encoding term element, the TEI syntax cannot readily convey that the LCSH string may have chronological or geographical components.[38] Therefore, because meaning is tied to the syntactic structure that expresses it, metadata transformations are semantic as well as syntactic operations.

So what is the benefit of discussing the core metadata-practice operations of mapping and transformation in terms of semantics and syntax? How do mapping and transformation relate to the previous discussions of metadata aggregates, layers of metadata responsibilities, and cross-community translations? The connections become apparent in light of the recommendation in the NISO *Framework of Guidance for Building Good Digital Collections* that digital collection developers create objects,

[36] Jane Hunter and Carl Lagoze, "Combining RDF and XML Schemas to Enhance Interoperability Between Metadata Application Profiles," under "2. Semantic Web Metadata Architecture" and "3. Combining RDF and XML Schemas." http://www.cs.cornell.edu/lagoze/papers/HunterLagozeWWW10.pdf (8 Mar. 2008); Carol Jean Godby, Devon Smith, and Eric Childress, "Two Paths to Interoperable Metadata," 2003, 2-3. http://www.oclc.org/research/publications/archive/2003/godby-dc2003.pdf (8 Mar. 2008).
[37] Godby et al., "Two Paths to Interoperable Metadata," 3.
[38] Text Encoding Initiative, "TEI Guidelines," http://www.tei-c.org/Guidelines/ (8 Mar. 2008).

metadata, and collections that are building blocks for reuse and integration.[39] By mapping and transforming metadata to enable the digital resources of diverse user communities to integrate successfully in information systems, metadata practitioners play key roles in facilitating multidisciplinary research and instruction. It is to these ends that metadata specialists create tools such as crosswalks and transformation scripts that support the semantic and syntactic interoperability of digital resources.

Metadata practitioners perform these operational tasks informed by research communities that are developing models and building services to facilitate the interoperation of digital resources created by and for diverse user groups.[40] Metadata practitioners design crosswalks and transformations, along with other metadata resources such as application profiles, XML namespaces, and XML schemas, to facilitate communication and cooperation across boundaries of communities of practice. Geoffrey Bowker and Susan Leigh Star, who conduct classification research at Santa Clara University and the University of California-San Diego respectively, call tools that perform such cross-community functions "boundary objects" because information specialists construct them at community boundaries in order to mediate the information needs of diverse communities.[41]

In arguments relevant to the present discussion, Bowker and Star recognize that the acts of creating boundary-spanning tools are dynamic processes in which

[39] NISO *Framework*, 1.
[40] Hunter and Lagoze, "Combining RDF and XML Schemas"; Paul Miller, "Interoperability: What Is It and Why Should I Want It?" *Ariadne* 24 (2000). http://www.ariadne.ac.uk/issue24/interoperability/ (8 Mar. 2008); William E. Moen, "Mapping the Interoperability Landscape for Networked Information Retrieval," Apr. 2001. http://www.unt.edu/wmoen/publications/MapInteropJCDLFinal.pdf (8 Mar. 2008); Friesen, "Semantic Interoperability"; Godby et al., "Two Paths to Interoperable Metadata"; Norm Friesen, "CanCore: Semantic Interoperability for Learning Object Metadata," in *Metadata in Practice*, 104-16.
[41] Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge: MIT Press, 1999), 15-16.

information practitioners actively engage with members of user communities.[42] As members of user communities help shape the tools that metadata practitioners create, so also will those tools shape the communities they serve. Over time, metadata tools—especially such high-visibility tools as controlled vocabularies and resource-type naming schemes—enter the work practices of the communities that adopt them. Moreover, as metadata practitioners build tools that translate information about digital resources from the language of one community to that of another, those tools become essential linkages in an infrastructure of resources, systems, and services that serve multiple communities and support inter-community communication.[43] By anticipating a path from the collection at hand to a more extensive digital infrastructure, metadata practitioners build semantic and syntactic links that span community boundaries to serve the broader interest of multidisciplinary collaboration.

Conclusion

The preceding observations have offered a multidimensional view of current metadata practice. As participants in digital-collection development efforts, metadata practitioners approach metadata in aggregates. Their informational role as investigators of user communities' communication practices informs both their interactions with their co-developers and their creation of metadata resources that address collection goals. One of their primary responsibilities on digital-collection development teams is to know enough about the communities of practice relevant to the collection to be able to

[42] Bowker and Star, *Sorting Things Out*, 254.
[43] Bowker and Star, *Sorting Things Out*, 286, 290-93, 296-98.

build metadata structures that mediate their linguistic and cultural differences. Moreover, in their operational contributions to digital collection building, metadata practitioners build tools that advance the cause of interoperability and cross-community communication.

These observations also hold strategic implications for library metadata operations. Because practitioners regularly inherit or harvest metadata en masse from heterogeneous community-based sources, wholly manual metadata processes do not scale.[44] This is consistent with the general migration in library technical services work away from data entry and toward data manipulation.[45] Rather than calling for record-by-record entry of metadata by specialists with similar skills, as in representative cataloging operations, current metadata efforts call for collaborations by practitioners with complementary skill sets. In CUL Metadata Services, for example, typical projects involve a metadata librarian to represent metadata interests on cross-community teams, to develop element sets, and to define content standards; a metadata technologist to create scripts for data conversion; and a metadata assistant to perform quality control checks and manual data clean-up. By recognizing the complementary metadata skills needed in digital collection efforts, libraries can benefit from organizing metadata operations that integrate automated and manual processes along these lines.[46] Indeed, strategic integration of automated and manual operations may be the essential niche for metadata practice in libraries.

[44] National Research Council, *LC21*, under "Metadata as a Cross-Community Activity."
[45] Calhoun, "Technology, Productivity and Change," 284.
[46] National Research Council, *LC21*, under "Metadata as a Cross-Community Activity."