

8th International Digital Curation Conference

January 2013

Toward Standardization: A Participatory Framework for the Process of Developing Scientific Metadata Standards

Lynn Yarmey,

National Snow and Ice Data Center
University of Colorado at Boulder

Karen S. Baker,

Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign

Abstract

In contemporary scientific research, standard-making and standardization are key processes for the sharing and reuse of data. The goals of this paper are twofold: 1) to stress that collaboration is crucial to standard-making and 2) to recommend recognition of metadata standardization as part of the scientific process. To achieve these goals, a participatory framework for developing and implementing scientific metadata standards is presented. We highlight the need for ongoing, open dialogue within and among research communities at multiple levels. Using the Long Term Ecological Research network adoption of the Ecological Metadata Language as a case example in the natural sciences, we illustrate how a participatory framework addresses the need for active coordination of the evolution of scientific metadata standards. The participatory framework is contrasted with a hierarchical framework to underscore how the development of scientific standards is a dynamic and continuing process. The roles played by 'best practices' and 'working standards' are identified in relation to the process of standardization.



Introduction

Expectations for the sharing and reuse of scientific data are increasing, as evidenced by current attention to data access technologies, funding agency data policies, and data management education. Though data in the natural sciences are heterogeneous and difficult to standardize (e.g. Madin et al., 2008; Reichman et al., 2011; Mayernik, 2011; Willis et al., 2012), data sharing and reuse depend, in part, on metadata standardization through the development and implementation of metadata standards. The development of scientific metadata standards is sometimes mistakenly perceived as a well-defined, one-time undertaking. While efficiencies-of-scale drive the process for developing standards in economic arenas, such frameworks do not adequately characterize the process of developing metadata standards for scientific data.

To open up discussion, we introduce a participatory framework for the development of science metadata standards grounded by respect for complexities-of-scale. A framework that is explicitly participatory 1) opens-up standards development efforts to ensure broader representation from science communities, 2) formalizes collaboration as a necessary component of standardizing scientific data descriptions, and 3) anticipates management of a trajectory of change throughout the process. We argue for envisioning standards development as an evolving, continuing design process as opposed to a set of defined steps leading to a solution in the form of an enduring standard. The term ‘framework’ is used instead of ‘standards lifecycle’ in order to emphasize relationships within broader, longer-term standardization efforts. Our aim is to foster dialogue and increase awareness of processes as central to developing useful and usable scientific standards.

The perspective of the authors is that of practicing information managers in the natural sciences. Years of experience with researchers, science networks, and data centers have shown how conceptual development and inclusive, collaborative efforts are essential to standardization.

Background

Standards are mechanisms of coordination that facilitate the exchange and comparability of information and products in the global community. Mature standards that are broadly accepted and implemented become authoritative resources for infrastructure development. As they exist across a number of arenas and carry a wide variety of functions, the study of standards is a multi-faceted endeavour. Previous work has addressed the classification of standards as well as the tension between standardization and flexibility (Sherif, 2001; Hanseth et al., 1996; Molka, 1992; Lehr, 1992, deVries, 2005). The development of all types of standards and the slowness of the process are topics of research (Bowker & Star, 1999; Lampland & Star, 2009; Jakobs, 2005, 2010; Lyytinen & King, 2006).

Standards were developed early on as physical specifications required for engineering and economic arenas – weights, measures, and strengths (Thompson & Taylor, 2008). In the current era of global interconnectivity, standards are needed for the exchange and coordination of digital information as well as of material goods. deVries (2005) assembled a diverse set of definitions that reveal the multiple roles for

standards as the need for standards beyond industry arenas was recognized (Rumble et al., 2005). The European Information and Communications Technologies Standards Board (ICTSB) defines a standard as a formal document within an existing context: a “document, established by consensus and approved by a recognized body, that provides, for common ... use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context.”

Hierarchical Framework for Development of Standards

Commercial and engineering standards play a significant role in successes and failures of systems and infrastructures (e.g. Edwards, 2003; Edwards et al., 2007; Star & Ruhleder, 1996). Schmidt and Werle (1998) report “The availability of standards not only helps firms to achieve economies of scale; it also facilitates the building of comprehensive, integrated technical systems.” The goal of commercial development is to create a final, fully-functional product quickly. Participants in this standard-creation process include large national or international bodies with economic and political perspectives. As in the case of typing keyboards (Lewin, 2002; David, 2007) and mobile communications standards (Akhtar, 2005; Fomin, 1999), the process frequently involves selection of one from existing options.

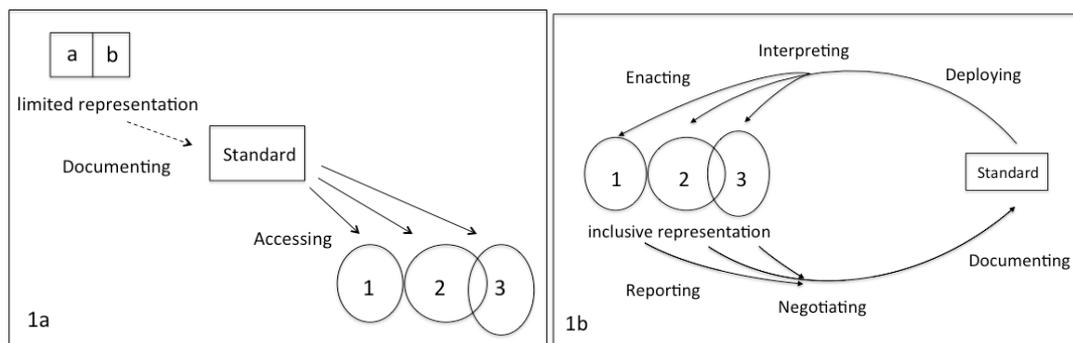


Figure 1. a) Hierarchical framework for developing standards where boxes a and b represent participants developing the standard; circles represent organizational units. b) Participatory framework with points of divergence (interpreting) and convergence (negotiating) shown as part of a cyclical process.

In a stylized presentation, Figure 1a presents a hierarchical framework for development of a standard using a stepwise ‘top-down’ approach. The upper boxes (labelled ‘a’ and ‘b’) signify exclusive agents who undertake creation of a standard through bounded, short-term discussions driven by efficiencies-of-scale. In this case, the process of creating standards is about reducing differences through elimination or selecting from alternatives, and/or through consolidation by generalizing details. The process is influenced by political considerations, limited time, and desire to maximize return on investment. Once a standard has been finalized, the development process is complete, the standard is published, and support for discussion ends. Individuals and groups (denoted 1, 2, 3) have access to the standard.

Metadata Standards in a Scientific Context

Shifting to a research context in the natural sciences, we find an arena of extreme complexity. To study the nonlinearities and interdependencies of interacting, living systems, scientific research incorporates rapidly changing observational methods that contribute to knowledge generation. The research process requires an adaptable infrastructure for organization of scientific data described by metadata (Greenberg et al., 2009; Greenberg, 2010). The National Academy of Science (NAS, 2009) defines metadata standards as descriptions of “the content, context, and structure of information objects, including research data, at any level of aggregation (for example, a single data item, many items, or an entire database).” In considering processes of collaboration and development for science metadata standards, we expand from technical factors and end products to socio-organizational factors and design as well.

The development of science metadata standards involves many different socio-organizational units, defined here as individuals or groups with their own situated culture, practices, resources, and needs. Examples of socio-organizational units include a researcher’s lab, a field project, a data center, a government agency, a field station, or a research community. Bringing these units together despite their differences is challenging, though there are a number of approaches that can help. *Design* is an integrative activity able to address complex issues as well as a multiplicity of organizational arrangements and political agencies (Green, 1996; Berg, 1998; Millerand et al., 2012). *Participatory design* highlights local engagement and expands to the long-term in *continuing design* (Schuler & Namioka, 1993; Dittrich et al., 2002; Karasti et al., 2010). Whelton and Ballard (2002) identify three design activity models: rational problem-solving, social process, and an experiential process providing a context for learning. Efforts in the field of information systems design have explored a variety of design approaches to support individual learning and group dialogue (e.g. Wagner et al., 2010; Millerand & Baker, 2010; Lyytinen, 1987). Design-oriented approaches for soft systems methodologies take broad views of human as well as technical requirements for information systems (e.g. Dix et al., 2004). Requiring human expertise in design helps ensuring that ‘bottom-up experience’ informs development of top-down, universal systems and standards (e.g. Hasselbring, 2000). In identifying standard-making as a design issue that crosses many socio-organizational levels, the dynamic interplay of social, political, and learning processes in the natural sciences is acknowledged and accommodated.

Heterogeneous data result from the many technologies, domains, and diverse sampling and processing methods used to study complex natural systems. Differences encountered in science are not necessarily seen as non-conformity or a failure to unify (Cragin et al., 2010; Star, 1991; Gasser, 1986). Heterogeneity is appreciated as that which informs scientists about systems at a variety of scales, though data reuse, automation, and even discovery are made extremely difficult by non-standard data and metadata. It is complexities-of-scale that define the scientific context though economies-of-scale, such as automated data processing and integration, are recognized as goals. Heterogeneous data and anomalies are features of the scientific process that require social and cultural as well as technical adaptation in standards development efforts. Scientific metadata standards should help document and mitigate heterogeneity. An existing multitude of standards however, has not been sufficient to achieve the promise of scientific data integration and reuse.

Metadata Standards in Practice

Metadata development initiatives have been underway for some time in national and international arenas. The Text Encoding Initiative (TEI) for marking up humanities documents was established in 1987, Dublin Core Metadata Initiative for cataloguing web-based resources started in 1995, and Darwin Core for documenting specimen data emerged around 1999. The Data Documentation Initiative (DDI) was funded in 1997 for describing social, behavioural, and economic science metadata. Within the observational sciences, the Federal Geographic Data Committee (FGDC) was established in 1993 and created a Content Standard for Digital Geospatial Metadata (CSDGM) for describing geospatial data. Subsequently, a paper by an Ecological Society of America committee was published on non-geospatial ecological data (Michener et al., 1997) from which developed the Ecological Metadata Language (EML). Though each standard has a story that provides insight into development efforts, our case draws on the EML story of scientific standards-making in practice.

The interpretation and enactment of EML by the Long-Term Ecological Research (LTER) program Information Managers for describing field ecology measurements provides an example of ongoing design activities. The LTER Network began in 1980 with a data manager included in each of six research teams. Today, the Network has twenty-six sites and an Information Management Community-of-Practice crossing disciplines as well as socio-organizational scales in a complex set of long-term arrangements. Continuity and oversight of EML are anchored at a technical center (Michener et al., 1997; Jones et al., 2006) while enactment activities are carried out at individual sites. EML was adopted by the Ecological Society of America and the Organization of Biological Field Stations; the LTER network adopted EML in 2001-2003 during the LTER's 'Decade of Synthesis' (Millerand & Bowker, 2009). The initial EML design process was open in theory but closed in practice given a 'fire-hose' of information precluding real-time participation by data providers; this resulted in a more a hierarchical approach to the development of the standard (Karasti et al., 2010, p395). As part of the enactment process, review criteria were developed against which site metadata quality and completeness could be judged (LTER IM, 2011, 2009). At the time however, important questions were only beginning to be formulated, i.e. 'what metadata is needed?' and 'what is a dataset?'. As understanding evolved, implementation participants realized that EML would require further refinement. Over time, LTER Information Managers iteratively worked with local scientists, shared site experiences and practices, and negotiated the standardization of content. Information Managers collaboratively developed a keyword controlled vocabulary (Porter, 2010) and a unit registry (Karasti et al., 2010) while scientists discussed language ambiguity and network-level developers pursued ontology-related data integration (Michener & Jones, 2012). As a case example, EML illustrates the interplay of hierarchical and participatory standard-making activities over time.

The initial development timeframe for EML was more than five years; this was followed by an even longer period of redesign, redeployment and re-enactment of the standard. The integrative work of identifying collaborative arenas, articulating requirements, and sharing, communicating, and negotiating across differing participant needs, traditions, and languages all takes time. The process of standardization, including the enactment of standards, is best recognized in the sciences as occurring at all times and at differing paces in various spheres-of-context

(Baker & Yarmey, 2009). As a result, the implementation of metadata standards is far less definitive and immediate than might be anticipated.

Challenges

Lack of adequate consideration for development frameworks makes it difficult to map expectations for science metadata standards to true standardization. Using a hierarchical development framework within a scientific context impacts timelines, roles, and expectations; the ramifications can be extensive. Introducing metadata requirements prematurely can lead to rushed and incomplete data representation, misguiding future efforts and defeating the objective of data reuse. Adoption of a ‘young’ metadata standard can stifle capture of contextual metadata at the local level, for example by missing new facets of methodology or integration-enabling elements as is currently the case with biological units. This can be perceived as prioritizing standards over scientific needs, exacerbating researcher disenfranchisement and widening the cultural gap between data professionals and scientists. Mismatching standards and functionalities is an issue as well; applying a metadata standard meant for discovery to capture deep contextual metadata may lead to inappropriate cross-walking of detailed data descriptions into general metadata fields. Further, high-level standards that generalize data description can increase data provider concerns about data misuse. Implementing a standard is further complicated in cases where the evolution of the standard is faster than enactment efforts, so that enactors are constantly playing catch-up. Adding urgency, the phenomena of path dependence (Puffert, 2010) describes the situation where an approach introduced early on may become ‘locked in’ regardless of its usefulness (Edwards et al., 2007; David, 2007).

Proposing a Participatory Framework

Considering the creation and implementation of standards as an activity completely separate from the conduct of science divides resources and increases the communication gap between metadata efforts and science. To increase bi-directional communication and engagement in scientific metadata work, we explore the characteristics of a participatory development process for scientific metadata standards (Fig. 1b) that more closely mirrors the conduct of science. Aligning with scientific research processes introduces an opportunity to partner with researchers in describing their science and data rather than imposing external metadata requirements. The proposed participatory framework has several key characteristics appropriate to the research context. First, the process is open and inclusive. Second the process is designed to maximize capture of scientific understanding rather than simply implement a standard. Third, scientific heterogeneity is consciously identified and strategically managed through informed consensus and negotiation. As with scientific work, the focus is not solely on final outcomes but on awareness and reconciliation of differences. Fourth, the process is ongoing, reflecting the long-term process of learning. The longer planning window allows flexibility for diverse participants to anticipate change and align as needed given local circumstances and resources.

Circles 1, 2, and 3 on the left of Figure 1b symbolize three different formal or informal socio-organizational units that could include scientists, librarians, standard-making bodies, and others. Each works within its own sphere-of-context characterized

by a unique configuration of organizational placement, project definition, technical aptitude, resources, research perspective, and available infrastructure. Rather than needing to have a specific skill set, participants with all types of expertise are welcomed and sought out. In addition to being participants in development of a scientific metadata standard, those engaged in the process are continuing learners. Standard-making participants co-construct an evolving scientific understanding by carrying out data curation in a research context.

The arrows on the lower left of Figure 1b depict the reporting or sharing of participants' practices, needs, and experiences with metadata (whether it is called 'metadata,' 'documentation,' or another term). Participants contribute expertise and lessons learned from their situated work. They may be in emergent roles such as data managers, data scientists, data librarians, etc. This early communication addresses the cultural need for relationship-building, increases participant buy-in, and offers each participant the understanding of future changes and process timelines. Formal and informal information exchange enables comparison across metadata collections and perspectives. Comparison leads to categorization. Differences exposed through reporting can be prioritized for classification and possible consolidation. Analysis of similarities and differences leads to new scientific understandings.

Moving along the lower part of Figure 1b, negotiation occurs at the point of convergence. Faced with disparate metadata, participants can decide upon reconciliation through an agreed-upon standard. Possible negotiation outcomes include resolution or postponement in addition to elimination and generalization. Ideally, negotiation is inclusive with each participant representing local priorities, infrastructures, resources, and needs. Negotiation points and agreements are articulated and captured through documentation. Documentation of scientific metadata standards and processes serves the same purpose as peer-reviewed publication; it keeps science open and cognizant of decisions that shape community understanding and research trajectories. As partners in the process, participants are empowered to take their new knowledge back to their local community for incorporation in decision-making and planning.

The outcome of the negotiation stage is a strategically released version of a standard (seen in the right-hand box). Necessary tools, guidelines, and other implementation mechanisms are identified and addressed through collaborative efforts. Once deployed (upper right arrow), the implementation process includes expected divergence as each socio-organizational unit separately interprets the standard in their respective sphere-of-context. Enactment (upper left arrows) involves essentially a site-specific implementation cycle of design, development, and deployment taking into account local circumstances (Millerand et al., 2012; Millerand & Bowker, 2009). Since the situated practices, systems, and infrastructures – human, cultural, and technological – differ, the standard and any new requirements must be aligned with existing practices. Any socio-organizational unit may find there are multiple ways to integrate the new standard with their own existing practices; formal or informal governance structures are used to decide among options. With standardization occurring in different ways depending upon local contingencies and readiness, timeframes and outcomes will differ as well.

Time may be needed for local redesign and to develop best practices for use of the standard. Best practices develop from local decisions made in interpreting a standard

to ensure that local circumstances are managed uniformly. A working standard, an ad hoc convention developed pragmatically to describe locally-developed procedures in response to a local need, may already exist and require modification or replacement. Alternatively, a site-specific working standard may develop as an enhancement to a community standard. As noted by Gasser (1986), local adjustments may appear initially as ‘out of specification’ but may in time contribute via the proposed participatory process.

Socio-organizational units with adequate support for information systems are in the unique position of developing site-specific metadata schemes that can be mapped to a variety of existing and emerging metadata standards. This enables site compliance with multiple standards while insulating local development efforts from dependence on external factors. Research-focused, local development efforts can adapt more rapidly than large-scale efforts, and, being closer to the science, can better capture requirements for data reuse. Sharing these situated experiences and understandings with others contributes to another round of the ongoing standard-making process.

Realities of Participatory Approaches

The participatory development of scientific metadata standards is distinguished from a hierarchical approach by the framework characteristics in Table 1. A hierarchical framework within a competitive market culture develops specifications in order to create an end-product rapidly. Alternately, a participatory framework within a research culture develops integrative mechanisms over the long-term that facilitate learning and support scientific inquiry.

Framework	Hierarchical	Participatory
Culture	Market	Research
Purpose	Specification	Integration mechanism
Goal	End product	Supporting ongoing scientific inquiry
Strategy	Competitive	Collective
Orientation	Solution	Staged cycles
Pace	Rapid	Slow
Influences	Politics, economics, technology	Existing knowledge, situated practices, technology
Dominant Driver	Economies-of-scale	Complexities-of-scale
Implementation Tools	Often existing	Often emerging
Participants	Limited	Inclusive
Standardization	Choosing between competing options	Coming to a shared understanding through learning

Table 1. Characteristics of two standard-development frameworks.

While participatory design can enhance thinking about standard-making, short funding cycles favor product-driven, hierarchical frameworks. Resources,

infrastructure, and expertise within the natural sciences are generally lacking as yet for inclusive activities, cooperation across socio-organizational units, and cross-community dialogue and engagement. Social scientists (e.g. Ribes & Baker, 2007), data curators (e.g. Cragin et al., 2010; Heidorn, 2011; Thompson, 2012), and data managers (e.g. Michener et al., 2011; Baker & Chandler, 2008) engage with scientists on these issues and contribute to the standards development process. It is also important to consider how the many scientific socio-organizational units engage with technical agencies and thereby shape subsequent requirements and policies.

The participatory process does not come without difficulties. It is time consuming, costly, and leans heavily on governance structures that often don't yet exist. Despite the difficulties, there are indications of potential success. In addition to the LTER EML case example, the field of health science has early experience in community-based participatory research (CBPR), an approach that recognizes and addresses the value of engaging with diverse public participants (e.g. Israel et al., 1998). Envisioning the metadata standardization process in scientific research using a participatory framework highlights the need for early planning to develop infrastructure that supports incremental changes throughout a staged process.

The Process of Standardization

Standardization refers to the process of aligning standards, practices, and content within and across socio-organizational units. Hierarchical standardization involves establishing compliance to set specifications while scientific metadata standardization requires a participatory strategy for growing a shared pool of knowledge. Standardization is essentially a process of learning; Ribes and Finholt (2007) note “the creation of a metadata standard could be viewed as a form of new knowledge.” The creation of science metadata standards developed in a participatory framework will necessarily involve negotiations to resolve disparate understandings and forge consensus through common terminology. The development process presents an opportunity to articulate and develop common views despite layers of complexity.

Collaboration is at the heart of human organization and growth of infrastructure (Lee et al., 2006; Bowker et al., 2010); it can take many forms depending on participant needs. While some commercial and scientific standards remove much of the need for interpersonal communication between those wishing to interoperate, the ongoing process of metadata standardization across the complex realms of science requires users and developers to work as co-designers in continuing collaboration.

Levels and Scales

Scientific research occurs in diverse socio-organizational units where the work of standardizing local practices is ongoing for each unit through a variety of activities. With both overlapping and contradictory interests, each unit has a sphere-of-context defined along local-remote, higher-lower, and/or large-small continuums that collectively may be called levels. *Level* refers to an ordering with differing scopes and responsibilities for the work that occurs in different contexts. From the perspective of an individual lab, a multi-investigator project is higher-level or more remote. From the perspective of a domain, however, a multi-investigator project may be at a lower level

and of narrower scope. Communication and standardization take place both across units at the same level and across differing levels.

Fomin (1999) argues that the standard is a boundary-crossing object that enables integration across different organizational levels. His multi-level analysis considers the roles and function of standards in terms of micro-meso-macro scaled categories associated with social, technical, and economic forces, respectively (Fomin, 2003). With scientific metadata standards, we map instead to three socio-organizational levels: local, community, and large-scale. Scientific metadata standards develop and are enacted at all levels, eventually interfacing with local, or micro, working standards and prompting development of best practices through the process of standardization. Community gateways develop at the meso-scale as data repositories such as the Global Biodiversity Information Facility (GBIF) and GenBank. At the macro-scale, recommendations about systemic solutions are meant to address large-scale issues.

Conclusions

This paper discusses standard-making as one part of the process of standardization for scientific metadata. A participatory framework is conceptualized and characterized by the expectation of ongoing collaboration. The framework highlights how standards themselves are only one aspect of a complex process; standards alone do not inherently achieve standardization. Data reuse and automated processing rely not on metadata shoehorned into ill-fitting standards to meet minimum requirements, but rather on the advancement of science through collaborative learning at many levels and in many spheres-of-context. In both the scientific arena and with metadata standards, a participatory approach is essential for supporting scientific inquiry.

The development of multiple standards is an important part of research and is by no means a failure to solve the problem of metadata development in the natural sciences. The cyclical standard-making process illustrates that divergence due to local interpretation and enactment is as important as the convergence that removes barriers to standardization. Both are integral to the evolution of scientific metadata standards and to scientific inquiry. Further study of standards enactment, the roles of best practices and working standards, and standardization across levels is needed.

Given the infinite variety of ways to organize and classify digital data that represent observations of our world, participatory development of scientific metadata standards is required to capture the complexity of natural systems. Bringing experts from different domains together through a participatory framework is a step towards supporting the research needed to answer grand challenge science questions of our day. In its provisionally, the process of standardization parallels the ongoing nature of the scientific knowledge-making process itself.

Acknowledgements

Support provided by Palmer Station LTER (NSF OPP-08-23101), California Current Ecosystem LTER (NSF/OCE-10-266607), and Data Curation Education in Research Centers (IMLS RE-02-10-0004-10), and the National Snow and Ice Data Center (NSF OPP-10-16048).

References

- [book] Akhtar, S. (2005). Pagani, Margherita (Ed.) *2G-5G Networks: Evolution of Technologies, Standards, and Deployment* (Second Ed.). (pp. 522-532). Hershey, PA: IGI Global.
- [journal article] Baker, K.S., & Chandler, C.L. (2008). Enabling long-term oceanographic research: Changing data practices, information management strategies and informatics. *Deep Sea Research II* 55(18-19): 2132-2142.
- [journal article] Baker, K.S., & Yarmey, L. (2009). Data stewardship: environmental data curation and a web-of-repositories. *International Journal of Digital Curation* 4(2), 12-27.
- [journal article] Berg, M. (1998). The Politics of Technology: On Bringing Social Theory into Technological Design. *Science, Technology & Human Values* 23(4), 456-490.
- [book] Bowker, G.C., & Star, S.L. (1999). *Sorting Things Out; Classification and Its Consequences*. Cambridge, MA: MIT Press.
- [book chapter] Bowker, G.C., Baker, K., Millerand, F., & Ribes, D. (2010). Toward information infrastructure studies: Ways of knowing in a networked environment. In J. Hunsinger (Ed.) *International Handbook of Internet Research* 97. Springer Science+Business Media.
- [journal article] Cragin, M.H., Palmer, C.L., Carlson, J.R., & Witt, M. (2010). Data sharing, small science and institutional repositories. *Philosophical Transactions of the Royal Society* 368, 4023-4038.
- [journal article] David, P.A. (2007). Path dependence – a foundational concept for historical social science. *Cliometrica – The Journal of Historical Economics and Econometric History* 1(2), 91-114.
- [book chapter] de Vries, H.J. (2005). IT Standards Typology. In Jakobs, K., *Advanced Topics in Information Technology Standards and Standardization Research*. (pp. 1-26). Idea Group Publishing.
- [proceedings] Dittrich, Y., Eriksén, S., & Hansson, C. (2002). PD in the Wild: Evolving Practices of Design in Use. In Binder, T., Gregory, J. & Wagner, I. (Eds.) *Proceedings of the Participatory Design Conference*. Malmö, Sweden, 23-25 June 2002. CPSR, Palo Alto, CA.
- [book] Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). *Human-Computer Interaction*. London: Prentice Hall.
- [book chapter] Edwards, P.N. (2003). Infrastructure and Modernity: Force, Time, and Social Organization in The History of Sociotechnical Systems. In T.J. Misa,

P. Brey, & A. Feenberg (Eds), *Modernity and Technology* (pp. 185-226).
Cambridge, MA: MIT Press.

- [report] Edwards, P.N., Jackson, S.J., Bowker, G.C., & Knobel, C. (2007).
Understanding Infrastructure: Dynamics, Tensions, and Design. NSF Report of a
Workshop on “History and Theory of Infrastructure: Lessons for New Scientific
Cyberinfrastructures”. Retrieved from <http://hdl.handle.net/2027.42/49353>.
- [proceedings] Fomin, V. V. (1999). The Dynamics of Standard Setting: Equilibrium
and Transformation in Social Networks. In K. Jakobs (Ed.), *Proceedings of First
IEEE Conference on Standardization and Innovation in Information Technology*,
Aachen, Germany.
- [journal article] Fomin, V.V. (2003). The Role of Standards in the Information
Infrastructure Development, Revisited. *Management Information Systems
Quarterly*, Special Issue: Standard Making: A Critical Research Frontier for
Information Systems, 147, 302-313.
- [journal article] Gasser, L. (1986). The Integration of computing and routine work.
ACM Transactions on Office Information Systems 4, 205-225.
- [journal article] Green, S. D. (1996). A metaphorical analysis of client organizations
and the briefing process. *Construction Management and Economics* 14, 155-164.
- [book chapter] Greenberg, J. (2010). Metadata and Digital Information. *Encyclopedia
of Library and Information Sciences*. New York, Taylor & Francis: 3610-3623.
- [journal article] Greenberg, J., White, H.C., Carrier, S., & Scherle, R. (2009). A
Metadata Best Practice for a Scientific Data Repository. *Journal of Library
Metadata* 9: 194-212.
- [journal article] Hanseth, O., Monteiro, E., & Hatling, M. (1996). Developing
information infrastructure: The tension between standardization and flexibility.
Science, Technology & Human Values 21(4), 407-426.
- [book chapter] Hasselbring, W. (2000). The Role of Standards for Interoperating
Information Systems, The Traditional Bottom-Up Integration. In K. Jacobs (Ed.),
Information Technology Standards and Standardization (pp. 116-130). Hershey,
PA: IGI.
- [journal article] Heidorn, P. B. (2011). The emerging role of libraries in data curation
and e-science. *Journal of Library Administration*, 51(7-8), 662-672.
- [online document] ICTSB, The Information and Communication Technologies
Standards Board. Retrieved from <http://www.icts.org/>
- [journal article] Israel, B.A., Schulz, A.J., Parker, E.A., & Becker, A.B. (1998).
Review of community-based research: Assessing partnership approaches to
improve public health. *Annual Review of Public Health* 19, 173-202.

-
- [book] Jakobs, K. (2005). *Advanced Topics in Information Technology Standards and Standardization Research*. New York: Idea Group Publishing.
- [book] Jakobs, K. (2010). *New Applications in IT Standards: Developments and Progress*. Hershey, PA: IGI Global.
- [journal article] Jones, M.B., Schildhauer, M.P., Reichman, O.J., & Bowers, S. (2006). The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution, and Systematics* 37, 519-544.
- [journal article] Karasti, H., Baker, K.S., & Millerand, F. (2010). Infrastructure Time: Long-Term Matters in Collaborative Development. *Computer Supported Cooperative Work - An International Journal* 19, 377-415.
- [book] Lampland, M., & Star, S.L. (Eds.) (2009). *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life*. Cornell University Press.
- [journal article] Lee, C.P., Dourish, P., & Mark, G. (2006). The Human Infrastructure of Cyberinfrastructure. *Computer Supported Cooperative Work Conference (CSCW)*, Banff.
- [journal article] Lehr, W. (1992). Standardization: Understanding the Process. *Journal of the American Society for Information Science* 43(8), 550-555.
- [book] Lewin, P. (2002). *The Economics of QWERTY: History, Theory and Policy*. New York: New York University Press.
- [online document] LTER IM (Information Management). (2011). EML Best Practices for LTER Sites. Retrieved from http://im.lternet.edu/sites/im.lternet.edu/files/emlbestpractices-2.0-FINAL-20110801_0.pdf
- [online document] LTER IM (Information Management). (2009). Review Criteria for LTER Information Management Systems. Retrieved from http://im.lternet.edu/sites/im.lternet.edu/files/LTER_IM_Review_Criteria_V1.1.pdf
- [journal article] Lyytinen, K. (1987). Different perspectives on information systems: problems and solutions. *ACM Computing Surveys (CSUR)* 19(1), 5-46.
- [journal article] Lyytinen, K., & King, J.L. (2006). Standard making: a critical research frontier for information systems research. *Management Information Systems Quarterly* 30(1): 405-411.
- [journal article] Madin, J.S., Bowers, S., Schildhauer, M.P., & Jones, M.B. (2008). Advancing ecological research with ontologies. *Trends in Ecology & Evolution* 23(3), 159-168.

- [thesis] Mayernik, M.S. (2011). Metadata realities for cyberinfrastructure: Data authors as metadata creators. *Information Studies*. Los Angeles, UCLA. PhD Thesis.
- [journal article] Michener, W.K., Brunt, J.W., Helly, J.J., Kirchner, T.B., & Stafford, S.G. (1997). Nongeospatial metadata for the ecological sciences. *Ecological Applications* 7, 330-342.
- [journal article] Michener, W.B., & Jones, M.B. (2012). Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology and Evolution* 27(1), 85-93.
- [journal article] Michener, W.K., Porter, J., Servilla, M. & Vanderbilt, K. (2011). Long term ecological research and information management. *Ecological Informatics* 6: 13-24.
- [journal article] Millerand, F., & Baker, K.S. (2010). Who are the users? From a single users to a web of users in the design of a working standard. *Information Systems Journal* 20, 137-161.
- [book chapter] Millerand, F., & Bowker, G.C. (2009). Metadata standards: Trajectories and enactment in the life of an ontology. In S.L. Star & M.Lampland (Eds.), *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life* (pp. 149-176) Cornell University Press.
- [journal article] Millerand, F., Ribes, D., Baker, K.S., & Bowker, G.C. (2012). Making an issue out of a standard: Storytelling practices in a scientific community. *Science, Technology, & Human Values* 37.
- [journal article] Molka, J. (1992). Surrounded by standards, there is a simpler view. *Journal of the American Society for Information Science* 43(8), 526-530.
- [report] NAS (2009). Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age. The National Academies Press, Washington, D.C.
- [newsletter] Porter, J.H. (2010). A controlled vocabulary for LTER datasets. LTER Databits Spring 2010. Retrieved from <http://databits.lternet.edu/spring-2010/controlled-vocabulary-lter-datasets>
- [online document] Puffert, D. (2010). Path dependence. *Economic History* EH.net. Retrieved from <http://eh.net/encyclopedia/article/puffert.path.dependence>
- [journal article] Reichman, O. J., Jones, M.B., & Schildhauer, M.P. (2011). Challenges and opportunities of open data in ecology. *Science* 703-705.
- [proceedings] Ribes, D., & Baker, K.S. (2007). Modes of social science engagement in community Infrastructure Design. In C. Steinfield, B. Pentland, M. S.

Ackerman and N. Contractor (Eds), *Proceedings of the Third Communities and Technologies Conference*, London: Springer.

- [proceedings] Ribes, D., & Finholt, T.A. (2007). Planning infrastructure for the long-term: Learning from cases in the natural sciences. *Proceedings of the Third International Conference on e-Social Science*, Ann Arbor, MI June 2006.
- [proceedings] Rumble, J., Carroll, B., Hodge, G., & Bartolo, L. (2005). Developing and Using Standards for Data and Information in Science and Technology. *Ensuring long-term preservation and adding value to scientific and technical data*. Edinburgh: The Royal Society.
- [book] Schmidt, S., & Werle, R. (1998). *Coordinating Technology: Studies in The International Standardization of Telecommunications*. Boston: MIT.
- [book] Schuler, D., & Namioka, A. (1993). *Participatory Design: Principles and Practices*. New Jersey: Lawrence Erlbaum Associates.
- [journal article] Sherif, M.H. (2001). A framework for standardization in the telecommunications and information technology. *IEEE Communications Magazine* 39(4), 94-100.
- [book chapter] Star, S. L. (1991) Power, Technologies and the phenomenology of conventions: On being allergic to onions. In J. Law (Ed), *A Sociology of Monsters: Essays on Power, Technology and Domination* (pp. 26-56). London: Routledge.
- [journal article] Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research* 7(1), 111–134.
- [report] Thompson, A., & Taylor, B.N. (2008). The NIST Guide for the Use of the International System of Units (SI), NIST Special Publication 811, 2008 Edition. Retrieved from <http://www.nist.gov/pml/pubs/sp811/index.cfm>
- [thesis] Thompson, C. A. (2012). *Answering the Call for Data Curation: An Exploration of the Careers of LIS Professionals Managing Data*. University of North Carolina, Chapel Hill. Masters Thesis.
- [book] Wagner, I., T. Bratteteig, & Stuedahl, D. (Eds.) (2010). *Exploring Digital Design: multi-disciplinary design practices*. Springer.
- [proceedings] Whelton, M., & Ballard, G. (2002). Wicked problems in project definition. *Proceedings of the Tenth International Group for Lean Construction*, Brazil.
- [journal article] Willis, C., Greenberg, J., & White, H. (2012). Analysis and synthesis of metadata goals for scientific data. *Journal of the American Society for Information Science and Technology* 63(8), 1505-1520.

