

Metadata Quality: From Evaluation to Augmentation

Diane I. Hillmann

Diane Hillmann is Research Librarian, Cornell University Library. She received a B.S. (1971) and M.L.S. (1976) from Syracuse University. She can be reached at dih1@cornell.edu or P.O. Box 81, Jacksonville, NY 14854.

Abstract

The conversation about metadata quality has developed slowly in libraries, hindered by unexamined assumptions about metadata carrying over from experience in the MARC environment. In the wider world, discussions about functionality must drive discussions about how quality might be determined and ensured. Because the quality-enforcing structures present in the MARC world—mature standards, common documentation, and bibliographic utilities—are lacking in the metadata world, metadata practitioners desiring to improve the quality of metadata used in their libraries must develop and proliferate their own processes of evaluation and transformation to support essential interoperability. In this article, the author endeavors to describe how those processes might be established and sustained to support metadata quality improvement.

Keywords

Metadata quality, metadata evaluation, metadata augmentation

Metadata Evaluation

As libraries re-examine their role in the information commons, their catalogs and their cataloging practices must also be re-examined. Not only must libraries serve their traditional role as organizers and providers of books and periodicals published by others, they must also seriously begin to take on the roles of publisher and aggregator of digital content. In these emerging bifurcated libraries of the future, many catalogers will find themselves operating outside the traditional boundaries of library cataloging, discovering the adventure of metadata. Although in many respects, catalogers are prepared and eager to move in this direction, they often find that their assumptions and preconceptions must be dusted off and re-evaluated.

In the traditional cataloging world, library management applications and bibliographic utilities (since the merger of OCLC and RLG, there is now effectively only one) have provided most of the quality control over metadata shared amongst libraries. Because of its continuing reliance on a “master record” concept, OCLC in particular developed extensive quality control mechanisms to ensure that the shared environment and the services they provide from it operate with as much quality as could be enforced, using primarily batch processes with some minimal human assistance. In this environment, with the utilities taking on most of the quality control efforts, libraries themselves have lagged behind in developing their own batch-oriented update and quality control

processes, basing most of their quality and updating programs on human review of individual records.

Given this background, it's not surprising that evaluation of metadata is not yet a routine part of the work of most metadata practitioners. Most library-based metadata professionals see their task as primarily one of metadata creation, and unless they inherit the care and feeding of metadata from other institutional projects, they tend to approach the problems of inadequate metadata quality from the viewpoint of a creator. Therefore, when in a situation arises where they have no control over the creation of metadata, they may find themselves frustrated by what they perceive as bad planning or judgment by the initial creator, and seek to solve the problems at the source rather than considering improvement strategies. Increasingly, however, they may find themselves managing data from multiple sources, aggregating that metadata to serve a particular purpose, often not the one for which the metadata was originally created. Because most metadata available for aggregation, whether within an institution or via harvest, was created in a context most likely rife with assumptions that it would be used narrowly and only in a specific context, any aggregation project automatically involves some confrontation with metadata quality issues.

Defining Metadata Quality

Conversations about metadata quality and quality improvement have been occurring for almost a decade, focusing primarily on determining how general quality criteria might be established in an environment characterized by increasing diversity of metadata formats and functions. Moen, Stewart and McClure, in their study of quality considerations for GILS metadata, described some of the important differences between traditional bibliographic control and metadata creation:

“Principles of bibliographic control certainly apply to the representation of networked resources in terms of rule-based creation (emphasis on structure and consistency to facilitate access), guidance by experts, and a consideration of user needs. But in practice, creation of metadata differs in several key respects. The resources to be described are volatile and distributed; no single, professional group has authority to dictate procedures; and not only are rules absent (at the Anglo-American Cataloguing Rules level of detail), there is no consensus that they should be created.” [1]

They go on to reflect on the diversity of resources and formats, which contributes mightily to the difficulty in designing evaluation processes with validity across the environment.

“Networked resources are highly heterogeneous, and various metadata schemes appear to reflect attributes assigned in a de facto fashion by different user communities. ... Given this force of user perspective on the representation of volatile information, and the lack of proven standards, systems of metadata ... may require uniquely tailored approaches to quality assessment. ... Schemes inevitably

represent a state of compromise among considerations of cost, efficiency, flexibility, completeness, and usability, and, thus, standards for quality must be based on the essential characteristics of each of these considerations. [2]

Given that there are unlikely to be fewer metadata formats in the future than there are at present, any solutions to the metadata quality problem must be designed to operate in an environment of considerable (and probably increasing) diversity.

In the MARC world, efforts like the Program for Cooperative Cataloging (PCC) developed detailed standards for record quality. In some cases, PCC standard records were less detailed than had been formerly expected in “full” records, but in other regards—for instance in mandating that all headings have authority records—the PCC record was more expensive to create, but it was hoped, more generally useful. Given that the PCC represented a relatively similar cohort of academic libraries, they represented the current community standard and best available thinking about what a quality record looked like and how it functioned.

In the metadata world, although some communities (most notably the Open Language Archives Community, known as OLAC) have coalesced sufficiently to agree on what quality meant in their context, for the most part communities using metadata are still floundering in their attempts to figure out where the best balance between “rich and comprehensive” and “efficient and functional” can be defined. Part of the challenge is that few communities of practice have been able to define their needs as a community and take the next steps to implement services that support their goals. OLAC very clearly saw that it was essential to build services that brought in the smallest and least well supported of their organizations, and so focused early on implementing their consensus about standards and extensions to Dublin Core using harvesting mechanisms to support community aggregation services. [3]

In contrast, libraries attempting to use Dublin Core or metadata standards other than MARC have been hampered by a view of their task much too tied to the MARC world. An example of this can be seen in some early criticism of the National Science Digital Library (NSDL). One of the persistent criticisms of early search results was that there were too many “duplicate records.” The notion of duplication as a quality problem is very much derived from the OCLC model, whose “master record” approach was based on the notion of libraries all using the same record for their cataloging. This made sense for OCLC—having libraries’ holdings spread between multiple cataloging records hindered the efficient use of the data for services like interlibrary loan—but it did not necessarily make sense to assume that the same model applied in the NSDL environment. The descriptive records contributed to OCLC were ostensibly created using the same cataloging rules, and according to cataloging convention should have been similar enough that distinctions between them might well be assumed to be the result of poor searching or less than optimal cataloging processes. In the NSDL world, descriptions of resources were usually based on a far more specialized view of the world, and were necessarily quite different. Thus the “duplication” was much less a quality problem than a display problem (in that the early NSDL had no way to avoid displaying all instances of

a resource description), and represented more an opportunity to re-think the notion of the “record” as the basis for providing search results to users rather than simply an indication of quality problems.

Clearly any useful discussion about quality in the metadata world requires at least that the quality standards not be assumed the same as those relevant in the MARC world, but instead be based on criteria more closely tied to the functionality sought for applications using metadata. Robertson, in his discussion of the impact of metadata requirements on LIS professionals, notes that this tie to functionality implies that there is no one answer to the quality question.

That different settings and purposes require different types of metadata quality should be no surprise as there are already other domains of knowledge management which have very different standards and purposes. The museum and the archive communities take a different approach to what represents quality in metadata. Museums record extensive detail about the provenance of an object as a necessary part of their purpose. Archives record extensive information but often only at the collection, rather than object, level due to the volume of materials they manage. These different purposes have existed side by side within the traditional knowledge management domain with little transference between. The metadata record for the same book will look very different in each setting and no one option is objectively better. [4]

Bruce and Hillmann, in their research into metadata quality, defined seven criteria and several levels of adherence, designed to re-start the quality discussion without those assumptions. As part of their seven criteria (completeness, accuracy, provenance, conformance to expectations, logical consistency and coherence, timeliness and accessibility) they note that differences in metadata might stem from either the environment in which the metadata is created and used or the differences inherent in the objects described. They also make the point that quality cannot be discussed in a vacuum: economic, political and technical constraints are a part of every decision affecting quality and perception of quality.

It is natural to ask which of the seven dimensions previously described is most important, or which most urgently needs to be present for a particular project. Where and how should we begin to foster quality? We believe that the way in which one might prioritize these various criteria is far from uniform, and are dictated by the nature of the objects to be described; whether the implementer is a source provider or an aggregator; and perhaps most importantly how the metadata is to be constructed or derived. Three familiar scenarios illustrate the diversity of options for metadata creation: a collection using author self-submission as the principle means of collecting both data and metadata; a project relying heavily on human judgment to create classificatory metadata; and a project using automated text-extraction techniques to pull metadata from a text corpus.

Each of these methods will have different ways of achieving high-quality results. Assuming a relatively stable corpus (a sometimes-dangerous but reasonable assumption), a computer program that extracts metadata will produce absolutely consistent results over an indefinite period of time, where a churning pool of student employees assigned to a markup project will not. A project where one person classifies information can make some assumptions about coherence and accuracy that a project relying on voluntary submissions cannot. A project that makes use of topical classification only as a means of creating rough boundaries for full-text search or a current-awareness service will not be as concerned about accurate classification as a project that is intended to produce a fine-grained taxonomic survey of a large body of literature. [5]

This introduction of the notion of context into the quality discussion has been a welcome development. Barton, Currier and Hey described the importance of context in their discussion of ePrint aggregation in the UK.

At a local level, the context in which the metadata is being created can have a bearing on the importance of quality assurance, particularly as it relates to specific fields. In some cases, the larger the dataset, the greater the likelihood that a problem will manifest itself. For example, in a large population of authors, name authority files may be needed to disambiguate one John Smith from another. In other cases, the degree of diversity can determine whether the quality of the metadata becomes an issue. For example, in an archive of papers and reports originating from a single research group, author affiliation can be set as a default value, whilst a subject-based archive may need to use corporate name authority files to ensure that papers can be retrieved effectively by organisation. However, in an environment where each repository or archive is part of a wider system predicated on interoperability, the importance of quality assurance for metadata creation goes far beyond that which the local context might suggest. The possible population of authors is that of the whole world, the diversity limitless. Metadata that supports successful resource discovery perfectly adequately in the local context may not be as effective in an aggregated system. [6]

Central to contextually oriented approaches is sufficient experience to determine the categories of problems—data that was missing, incorrect, confusing or insufficient—as defined by functional requirements within the relevant environment. In NSDL the diversity of resources made the definition of problem issues quite different from those in the ePrints environment, for example, where the resources described were very similar. Differing creation methods also had an impact: most of the ePrint projects described by Barton's group used author submission as their creation method, while most NSDL projects were crosswalking their data from internal databases into the Dublin Core format that NSDL preferred.

This necessity to keep in mind the required functionality, the nature of the described resources, as well as the metadata format conveying the information, contributes

significantly to the challenge of determining the appropriate level of quality for the task, as well as the effort required to evaluate the metadata for those qualities.

Evaluation Techniques

In the traditional library world, where most metadata has encountered at some point the bibliographic utility quality control algorithms and software applications' edit checks, most problems were either typographical errors or outdated headings. More serious quality surprises were few in this environment and most likely attributable to inadequate training and supervision. Thus, managers of MARC databases rely heavily on each other to maintain a reasonable level of predictability, and other than generalized sweeps for common typos, few batch oriented quality control strategies exist in traditional libraries exchanging records within the MARC record infrastructure. On occasions when records are imported from outside this environment, random sampling for quality evaluation is the norm, though ordinarily library database managers do some testing of the imported file in a non-production context. Since most batches of records are acquired from well-known suppliers, there is little impetus to invest in other methods to assess record quality and adherence to standards.

Attempts at describing metadata evaluation techniques appropriate to in this environment have been few. Dushay and Hillmann, in their paper detailing the evolving understanding of evaluation in the NSDL, describe moving from random sampling to more batch evaluation techniques using spreadsheets and specialized software for graphical analysis. [7] This strategy focuses on the characteristics of the "forest," which assists in identifying the anomalous "trees" as well as the patterns common to the set. In an aggregation setting, like the NSDL, quality problems are endemic, as each provider operates in an environment where local conditions determine how metadata issues are resolved and more general issues are poorly understood.

The California Digital Library (CDL) has also begun to develop evaluation strategies, with goals similar to those of NSDL, as a necessary precursor to any attempt at transformation.

"Metadata analysis should be able to answer a number of important questions, for example:

- * Which metadata fields are present?
- * What percentage of the total number of records have each field?
- * How consistent is the metadata within those fields?
- * What patterns can be detected?" [8]

Clearly, then, it is the aggregators where the early work is being done, but as individual libraries build their digital library programs, the techniques developed by these aggregators will enable evaluation processes within institutions to evolve, and important quality improvement services to be built and thrive.

Improving Metadata Quality

Identifying metadata problems is only the beginning: more critical is the task of ensuring that the effect of those known problems will not prevent users from retrieving the resources they need. Most of the focus of those addressing the metadata quality problem has been at the metadata creation end. Barton, et al. make that case firmly:

There will always be some aspects of the metadata that are inaccurate, inconsistent or out of date, even in systems which have extensive quality assurance procedures in place and have invested heavily in the creation of good quality metadata. For example, when a published subject classification scheme is updated, new resources may be classified using new subject terms but existing resources may not be reclassified, giving rise to inconsistent subject-based searches. Furthermore, in established systems, there may be a drift over time between policy and practice; a study into cataloguing practices in Scottish libraries as part of the CAIRNS Project found this issue to be widespread. Nevertheless, it is essential that quality assurance is built into the metadata creation process at the outset, that its scope extends beyond the local context and that the resulting metadata is as 'good' as it can be within the inevitable limitations of time and cost. [9]

Similarly, Guy, Powell and Day, in their discussion of quality improvement in ePrint archives, stress processes and feedback loops designed to improve metadata creation tools and documentation, rather than developing strategies for improvement without reference to initial providers of metadata.

It is important that whatever constructive feedback is obtained from testing be passed back into the system through redesigns of the application profile, controlled vocabularies, cataloguing guidelines and data entry tools. This QA [quality assurance] process can be viewed as a cycle or feedback loop, in that each stage feeds into the next ... Processes, standards, tools and documentation are iteratively enhanced with the overall aim of improving the metadata created. When good Quality Assurance is implemented there should be improvement in the quality of the metadata, the usability and performance of the eprint archive and there should be a decreasing rate of defects. [10]

In a single library or relatively limited aggregation environment, focus on source quality may well be a reasonable place to begin, but as metadata distribution becomes more widespread, and resources aggregated more diverse, focus on initial quality cannot be sufficient. Particularly in a world where metadata harvesting occurs from providers with no real knowledge of metadata issues and few incentives to spend their resources on improvements that don't benefit them directly, a focus on feedback to providers is of limited usefulness.

Clearly, metadata practitioners are unlikely to be rescued by any metadata analog of the MARC-based bibliographic utilities as they search for solutions to the quality problems.

Similarly, they are unlikely to find themselves working within a unified system of metadata standards such as libraries enjoyed during the heyday of MARC. This reality can be seen as both challenge and opportunity, and some possible strategies that do not focus only on metadata creation have been surfacing from the larger aggregators.

Where there is little hope of effecting improvement at the provider end, effort of necessity moves towards improving the data at hand and making those improvements available to others. NSDL, while spending significant time attempting to educate providers about quality metadata, also developed some of the earliest metadata normalization routines under the rubric of “safe transforms,” where some general strategies could be applied to all record instances without fear of degradation of the data. [11] The CDL has built upon some of these normalization ideas as well, and both aggregators identified as well the need to identify distinct batches of data where more specific transformations could be used.

As mentioned earlier, because NSDL was harvesting a broad range of science and math oriented data from a large number of providers, it early recognized that those projects recommending high value resources were choosing to describe some of the same resources, with differing emphases based on their topical specialty. Thus the notion of an aggregation basing its services primarily on “records” was quickly perceived as inadequate, and that view prompted an approach based instead on the idea that the “statements” in records could be aggregated into “descriptions” that were a fuller and more useful than the original separate records themselves. [12] This approach differs significantly from those developed in the MARC environment, where the desire for a single best record required that “duplicates” be rejected, and the decision was most often based on the source of the record or which one came first rather than any extensive evaluation of the records themselves.

Clearly, any strategy based on harvesting, where updates to individual records must be accommodated and any normalization applied by an aggregator preserved as well, had to be able to keep straight the source of any particular statement, in order to avoid having processes stepping on one another. This is particularly important where automated scheduling drives the sequence of processes, as Phipps, et al. have described. [13]

Concerns about the sequence of processes is relevant whether the process was *transformative*, e.g., based on modifying metadata based on the structure or values already available in statements, or whether it is *augmentative*, e.g., adding information based on information gleaned from the resource itself or some process upon that resource. These distinctions between transformation and augmentation are especially important when considering relevance and rating of multiple statements. In this view, transformations do things like:

- Detect controlled vocabulary values and attribute those values to a particular vocabulary
- Detect and fix common typographical errors
- Deprecate "promiscuous defaults," e.g., values that provide no information value, added to metadata only to fill a slot or provide functionality only

within a particular context, with no relevance outside that context

Augmentation, on the other hand, operates by using a variety of methods to access the described resource and add new value to the aggregated description. Some examples of this could be:

- Machine-based processes that add values, for example, topics or formats
- Human-based augmentation, such as the addition of:
 - Topics (whether expert-added or edited from machine assignment)
 - Relationships to educational standards
 - Other controlled vocabularies (for genre, etc.)

In both cases, proper sourcing of the statements, and in the case of normalization, creating separate statements rather than substituting new statements for inadequate ones, are critical to the ability to generalize and scale these strategies. A description consisting of aggregated sourced statements is susceptible to a variety of processes designed to provide downstream users with configurable descriptions based on their needs and capabilities. Over time, multiple statements can be rated using various criteria, and only the “best” used for exposure to downstream users who would rather not do the rating themselves. This strategy provides a very robust basis for metadata improvement services, and in addition integrates well with semantic web compatible methods for information maintenance using minimal human intervention. [14]

Terminology services

One example of quality improvement is terminology services. The recent UK Joint Information Services Committee (JISC) report on these services provides a useful definition:

“Terminology Services (TS) are a set of services that present and apply vocabularies, both controlled and uncontrolled, including their member terms, concepts and relationships. This is done for purposes of searching, browsing, discovery, translation, mapping, semantic reasoning, subject indexing and classification, harvesting, alerting etc. They can be m2m or interactive, user-facing services and can be applied at all stages of the retrieval process.

TS can be confusing in that they span very different application areas, vocabularies, communities, and can provide quite different kinds of services. They can be applied as immediate elements of the end-user interface (e.g. pick lists, browsers or navigation menus, search options) or can underpin services behind the scenes.” [15]

Terminology services coupled with normalization provide powerful tools for metadata improvement. As an example, consider a situation where a metadata provider uses a topical vocabulary for their metadata, but does not express the fact that the topical values within their metadata records come from this vocabulary—they may instead express this in their OAI description of the set, in their associated documentation, or in some other indirect manner. As part of a normalization routine for this particular set of records, a

service could encode the statements so that the values were properly designated as coming from that specific vocabulary. As an additional step, a terminology service could access a file or registry containing this vocabulary, and add or substitute the URI for that value (if there is one), update the preferred term if it changed since the term was originally applied, and/or cache or harvest the structure of the term for use within an application. Another service might then map the provider-supplied values to other relevant vocabularies, providing additional topical richness with little added investment of scarce human resources.

It seems likely that as the mass digitization efforts proceed, with their promise to provide scanned and searchable versions of the scholarly record, that the sheer scale of the effort and the age of many of the items will provide a fruitful opportunity to explore how metadata and full text can be used together to improve user experiences. For example, a century old treatise on medicine is likely to use outdated terminology to describe diseases and medical techniques, even as the metadata (if MARC with MeSH headings) will reflect more current topical terms. A user who knows the current terms, but not the historical ones, is likely to miss the treatise if searching only full text, but with an intelligent application making the connection between the terms using the existing MeSH term relationship structure, would be more likely to find the treatise.

Phipps, Hillmann and Paynter suggest that such services could be orchestrated without significant human intervention, maximizing the value of multiple services by intelligent scheduling.

“Using the term ‘orchestra’ to describe the array of services we envision emphasizes two important characteristics of this approach. First, the services we describe below are external to the harvesting task; each a separate player, specialized to perform particular operations, who can be called upon to provide specific outputs. Second, these complex parts are coordinated by a central intelligence—a ‘conductor’ who is not a player, but whose contribution is vital to the performance as a whole. Each of the services has their own roles, characteristics and relationships with other players—they resemble a group of musicians waiting for their cues.” [16]

Adherence to Standards

If “adherence to standards” is still the mantra for libraries, then what are the standards to which we should adhere? This question is hardly rhetorical, but the answer is certainly less simple than the one libraries have been answering so firmly in the MARC world. Most of the standards we might refer to in our attempt to be conformant in the metadata environment are still to some extent in development, and none are or will be as ubiquitous as MARC, or as broad ranging. For instance, the MARC standard includes not only the encoding for descriptive records, but also the transfer protocol used to share them. In the metadata world, standards tend to have a much smaller footprint, and in some sense must be assembled to meet specific goals and conditions, rather than just adhered to.

Given the diversity of use environments, resources, and desired functionality, the critical piece is not necessarily obeisance to “standards” but instead making intent transparent—documenting what we do and have done to create and transform, so that others can determine the suitability for re-use in their environment. In a shared metadata environment, it is impossible to distinguish quality metadata from marginal metadata, without reference and access to extensive documentation specific to the particular metadata being shared. Most institutions using OAI-PMH will be distributing metadata in a number of formats, and the essential tasks of documenting that metadata might begin with answering essentially the same questions journalists are trained to address:

- Who created the data (machines, humans, both?)
- Where was the metadata created (what is the institutional genealogy)?
- What is the metadata describing?
- When was it created, updated, etc.?

Bruce and Hillmann define three levels of quality, based partially on the extent of documentation provided to downstream users, whether they be partners, aggregators, applications, or users. [13] Application profiles, which at the most basic level document the intent of the creator of the metadata, give important clues to those outside the institution or domain of the metadata creators and are increasingly used to provide guidance to specific organizations and communities of practice. Eventually, machine-readable application profiles will allow machines to evaluate and validate metadata based on the expectations defined in the AP.

Conclusion

Important developments in the infrastructure required to support creation and maintenance of quality metadata are in process but not complete, making the challenge considerable for practitioners. In the meantime, it’s critical that more of the attention now clustered at the creation end of the discussion begin to shift towards the development of strategies to improve metadata as it moves from its source to different contexts of re-use. Developing a working metadata improvement environment depends on concerted efforts by those who can envision, build and manage these services. As part of this effort they must develop a way for all to participate as well as build sustainable methods to enable the services to survive and grow.

References

[1, 2] Moen, William E., Erin L. Stewart, and Charles R. McClure. “The Role of Content Analysis in Evaluating Metadata for the U.S. Government Information Locator Service (GILS): Results from an Exploratory Study, 1997.”

<http://www.unt.edu/wmoen/publications/GILSMDCContentAnalysis.htm>

[3] Bird, Steven, and Gary Simons. “Building an Open Language Archives Community on the DC Foundation.” In Hillmann, D., ed. *Metadata in Practice*, American Library

Association, Chicago, 2004. <http://www ldc.upenn.edu/sb/home/papers/mip.pdf>

[4] Robertson, R. John. "Metadata quality: implications for library and information science professionals," *Library Review*, 2005, v. 54, no. 5, pp. 295-300.

[5] Bruce, Thomas R. and Diane I. Hillmann. "The Continuum of Metadata Quality: Defining, Expressing, Exploiting." In Hillmann, D., ed. *Metadata in Practice*, American Library Association, Chicago, 2004. <http://hdl.handle.net/1813/7895>

[6] Barton, Jane, Sarah Currier, and Jessie Hey. "Building Quality Assurance into Metadata Creation: An Analysis based on the Learning Objects and E-Prints Communities of Practice." Paper presented at DC-2003, October 1003, Seattle, WA. <http://www.siderean.com/dc2003/201-paper60.pdf>

[7] Dushay, Naomi, and Diane I. Hillmann. "Analyzing Metadata for Effective Use and Re-Use." Paper presented at DC-2003, October 2003, Seattle, WA. <http://hdl.handle.net/1813/7896>

[8] Tennant, Roy. "Specifications for Metadata Processing Tools." 2007, 1/2, California Digital Library, 2004. http://www.cdlib.org/inside/projects/harvesting/metadata_tools.htm

[9] Barton, op cit.

[10] Guy, Marieke, Andy Powell, and Michael Day. "Improving the Quality of Metadata in Eprint Archives." *Ariadne*, 2004, Issue 38. <http://www.ariadne.ac.uk/issue38/guy/>

[11] Hillmann, Diane I., Naomi Dushay, and Jon Phipps. "Improving Metadata Quality: Augmentation and Recombination." Paper presented at DC-2004, Shanghai, China. <http://hdl.handle.net/1813/7897>

[12] Hillmann, op cit.

[13] Phipps, Jon, Diane I. Hillmann, and Gordon Paynter. "Orchestrating Metadata Enhancement Services: Introducing Lenny." Paper presented at DC-2005, September 2005, Leganés (Madrid), Spain. <http://arxiv.org/abs/cs.DL/0501083>

[14] Phipps, op cit.

[15] Tudhope, Douglas, Traugott Koch, and Rachel Heery. Terminology Services and Technology: JISC State of the Art Review. Joint Information Systems Committee (JISC), Bath, Eng., 2006. http://www.jisc.ac.uk/media/documents/programmes/capital/terminology_services_and_technology_review_sep_06.pdf

[16] Phipps, op cit.

[17] Bruce, op cit.