Metadata Quality, Utility and the Semantic Web: The Case of Learning Resources and Achievement Standards

Stuart A. Sutton

ABSTRACT. This article explores metadata quality issues in the creation and encoding of mappings or correlations of educational resources to K-12 achievement standards and the deployment of the metadata generated on the Semantic Web. The discussion is framed in terms of quality indicia derived from empirical studies of metadata in the Web environment. A number of forces at work in determining the quality of correlations metadata are examined including the nature of the emerging Semantic Web metadata ecosystem itself, the reliance on string values in metadata to identify achievement standards, the growing complexity of the standards environment, and the misalignment in terms of granularity between resource and declared objectives.

KEYWORDS. Semantic Web, metadata quality, achievement standards, K-12 educational resources

---

Stuart A. Sutton, JD, LLM, MLIS, PhD, is Associate Professor, The Information School, University of Washington. He serves on the DCMI Advisory and Usage Boards and served as Co-Chair of the DCMI Education Working Group from 1999-2006. He also serves as technical lead for the NSF-funded Achievement Standards Network.

The research upon which this article is based is partially funded by the National Science Foundation under Grant Nos. DUE-0121717 and DUE-0532854.

Cataloging & Classification Quarterly, Vol. 46(1) 2008
Available online at http://ccq.haworthpress.com

doi:10.1080/01639370802183065

INTRODUCTION

The correlation or mapping of learning resources such as lesson plans, curriculum units, and learning objects to formally promulgated achievement standards is a growing imperative in the K-12 environment. What we call achievement standards in this article are frequently called curriculum objectives in the cataloging literature and academic standards, curriculum standards, learning indicators and other names by the K-12 community. We choose achievement standards as a generic term indicating all forms of statements defining what K-12 students should know or be able to do as formally promulgated by a community to help shape teaching and learning activities in schools.

Starting slowly with the clarion call of *A Nation at Risk: The Imperative for Educational Reform*,[1] development of policies defining accountability for U.S. teachers and schools has accelerated the processes of standards-based education in the U.S. Largely unheard of in the U.S. at the beginning of the 1990s, every state in the Union except one has promulgated achievement standards defining what K-12 students will learn, when that learning will take place, and how learning will be assessed. Influences such as the federal No Child Left Behind Act of 2001, testing regimes such as the National Assessment of Educational Progress (NAEP) and state high-stakes testing are major drivers in the developing call for learning resources that assist teachers in meeting the demands of demonstrable accountability lurking behind the articulated state standards.

This article explores issues of quality in the creation and encoding of correlations in metadata. In particular, we frame the metadata quality issue in terms of the emerging Semantic Web and its enabling technologies. We posit that there are a number of forces at work in determining the quality of correlations metadata in the context of the Web–the nature of Semantic Web metadata ecosystem itself, the reliance on string values in metadata to identify achievement standards, the growing complexity of the standards environment, and the misalignment of resource and declared objectives granularity given current structural limitations. It is the goal of this article to identify the nature of these forces and to suggest solutions to address them.

We note initially that the notion of providing achievement standard information in metadata describing learning resources is a relatively new phenomenon. The process of recording this sort of information in bibliographic records has only been a formal part of the MARC world since the creation of Curriculum-Enhanced MARC (CEMARK) in

1993.[2] The goal at that time was for MARC to better serve the K-12 environment by providing the means for school librarians and other catalogers to make statements about resources as they pertain to the K-12 audience and curriculum. CEMARK achieved this goal through: (1) revision of a number of existing tags; and (2) the creation of tag 658 (Curriculum Objective) to record learning outcomes defined at national, state and district levels.

Since 1993, the MARC tag 658 has provided a limited means for capturing main and subordinate learning objectives. The tag includes a subfield code for noting what MARC calls a "Correlation Factor"–an uncontrolled value string identifying the degree of fit between the resource being cataloged and the learning outcome. While it is not our intention to discuss the metadata quality issues in the context of the MARC record, we note that a number of the factors affecting metadata quality assessments in the Web environment substantially impact MARC-based correlations.

In the following sections of this paper, we will first address what it means to say that a unit of metadata is of good quality–whether that unit is a single metadata statement or what we traditionally call a complete metadata "record." In essence, we will challenge the functional characterization of some of the quality indicators described in the research literature as denoting the "goodness" or "badness" of metadata when viewed in terms of the emerging Semantic Web metadata ecosystem. In the course of that discussion, we will draw what we believe is a useful distinction between metadata quality and metadata utility. We will then discuss the emerging ecosystem in which the metadata of concern to our inquiry resides and interacts. Having provided a framework for metadata quality and utility evaluation in the context of the Semantic Web, we will close by addressing these issues in achievement standards correlations in that environment.

METADATA QUALITY VS. UTILITY

Concerns over the quality of metadata date back at least as far as the emergence of standards for creating bibliographic records including the "library hand." Today, a number of empirical studies have explored the various dimensions of Web-based metadata quality including Moen, Stewart, and McClure,[3] Bruce and Hillmann,[4] Tennant,[5] Barton, Currier, and Hey,[6] Stvilia, Gasser, Twidale, Shreeves and Cole,[7] Johanis,[8] and Guy, Powell and Day.[9] Most of these studies discuss quality in terms of metadata generation while a few discuss it in terms of post harvesting refinement and augmentation.

Many of the studies define various quality dimensions or characteristics of quality metadata. Stvilia et al. assert that "[a]lmost as many different taxonomies of [information quality] dimensions have been proposed as there are writings about [information quality]."[10] From their examination of previous research, Gasser and Stvilia derived a framework consisting of thirty-eight dimensions divided into three categories–(1) intrinsic information quality; (2) relational/contextual information quality; and (3) reputational information quality.[11]

Bruce and Hillmann ferret out a useful set of quality characteristics that could well be framed as dimensions within Gasser and Stvilia's three categories. Building on the work of Johanis,[12] Bruce and Hillmann identify the following characteristics of quality metadata:[13]

- *Completeness*. The metadata schema (attribute space) used is capable of describing the intended resources as completely as necessary and possible; and, the elements in the schema are used in description as comprehensively and consistently as feasible.
- *Accuracy*. The values provided in the metadata, among other things: (a) are "correct and factual"; (b) are free of typographical errors; (c) use standard abbreviations; and (d) are conformant in the use of personal and place names.
- *Provenance*. The agents (whether human or machine) responsible for the creation of the metadata are identified as are the methods and controlling standards applied in the processes of creation and transformation.
- *Conformance to expectations*. The metadata schema and schemes (value spaces) should contain *all* and *only* the elements and vocabulary terms needed to support the defined purposes of the community to be served.
- *Logical consistency and coherence*. The metadata schema and schemes are used in a way that is consistent across the collection of resources and within any related discourse or practice communities.
- *Timeliness*. The metadata accurately reflects the current state of the resource being described.
- *Accessibility*. The metadata presents few or no barriers to physical and intellectual access to its content.

The characteristics identified by Bruce and Hillmann are not mutually exclusive. Citing Moen et al.,[14] Bruce and Hillmann note that it is difficult to determine whether the absence of a metadata element in a record is a *conformance to expectations* problem or a *completeness* problem when that record is viewed by a community of discourse or practice other than the community for which the record was created. In fact, we might posit that a metadata record that *conforms to expectations* of one discourse or practice community is absolutely bound in terms of *completeness*, *conforms to expectations*, and intellectual *accessibility* to be deemed of lesser quality in a different community. Thus, many of the quality criteria as defined in these various studies are highly conditioned by context–both the context for which the metadata was created and the context in which it may later find itself.

Guy, Powell and Day assert that the base metric in measuring the quality of metadata is its "fitness for purpose."[15] Since different communities of practice have different purposes, there is a relatively common assertion echoed in empirical studies of metadata quality that the context of use by a community of practice is a critical factor in determining the quality of metadata describing the resource. "Studies have repeatedly shown that information quality assessments are contextual."[16]

For example, Bruce and Hillmann identify *conformance to expectations* as an important quality criteria and frame it in the following contextual terms:

> Standard metadata element sets and application profiles that use them are promises from the metadata provider to the user. Moreover, they are promises surrounded by the expectations of the community about what such promises mean, how realistic they are, and how they are to be carried out. . . . Finally, metadata choices need to reflect community thinking and expectations about necessary compromises in implementation.[17]

Nilsson, Palmér, and Naeve note that "a single piece of media like a photograph can have a different meaning when used in a History context than when used in a Photography context."[18] While there may be substantial overlap in attributes of interest to both communities, it is inevitable that there will be any number of metadata statements about attributes critical to effective use of the resource in either community that are not shared by the other. As a result, both the history and the photography communities will view the metadata records fit for use in the other community as incomplete and non-conforming in the sense that "[element sets and application profiles] should not contain false promises, i.e., elements that are not likely to be used because they are superfluous, irrelevant, or impossible to implement."[19]

Thus, we are faced with a paradox: The quality of all metadata is both good and bad when assessed by various communities of practice through lenses framed by their community discourse, customs and practices. There appears to be no way to win the metadata quality game without addressing directly the source of the paradox. We posit that the fuzzy nature of the problem stems from the conflation of two related notions–quality and utility–into our current conception of quality. We posit further that if we separate these two notions, we can then assert of any given unit of metadata that while it may be of invariant good quality, its utility varies across knowledge and practice domains.

We can attempt such a quality/utility sort of Bruce and Hillmann's quality characteristics into: (1) those characteristics applicable as quality indicators within the context of a specific community of practice whose perceived quality shifts when the community context shifts; and (2) those characteristics whose perceived quality is contextually invariant. Such a sort is at best "rough" since the semantics of the current characteristics are strongly influenced by their characterization as "quality indicators."

In general, Bruce and Hillmann's *accuracy* characteristic and the *intrinsic quality indicator* category defined by Stvilia et al. fall within a single conceptual space. They can be framed as contextually independent and may be properly characterized in all instances in terms of quality with its connotations of "good" and "bad."

> Some dimensions of information quality can be assessed by measuring attributes of information items themselves, in relation to a reference standard. Examples include spelling mistakes (dictionary), conformance to formatting or representational standards (HTML validation), and information currency (age with respect to a standard index date, e.g., "today"). In general, intrinsic [information quality] attributes persist and depend little on context, hence can be measured more or less objectively.[20]

In Table 1, we set out our rough sort of Bruce and Hillmann's characteristics in terms of their contextual dependence (utility) or independence (quality).

The characteristics considered contextually dependent were deemed "quality issues" *solely within their initially intended context* and "utility issues" outside of it. The characteristics considered contextually inde-

TABLE 1. Bruce and Hillmann's Quality Characteristics as Quality and Utility Factors

Extrinsic Characteristics (Utility)	Intrinsic Characteristics (Quality)
Completeness	Accuracy
Accessibility (intellectual)	Accessibility (physical)
Conformance to expectations	Timeliness
Consistency and coherence	Provenance

pendent were deemed "quality issues" in all instances. Thus, *completeness* is a quality issue that can be judged solely within the initially intended context. A complete unit of metadata in its native context may present utility issues in other contexts where the absence of a particular piece of metadata lowers its fitness for purpose. Depending on the full scope of their semantics, the same conclusion can be reached with physical *accessibility* and *conformance to expectations*–they are quality issues solely within their initially intended context and utility issues outside of it.

The same cannot be said for *accuracy*. As noted by Stvilia et al., accuracy is an intrinsic attribute of the data and can be judged independently of context of use. Whether the unit of metadata is *consistent and coherent*, contains accurate *provenance* information, is physically *accessible*, or has problems with *timeliness* are quality judgments independent of context.

We consider this distinction between quality and utility useful when we shift our perspective to the emerging environment of metadata on the Web, where, as we shall see, the context of any single unit of metadata is malleable.

THE EMERGING WEB-BASED METADATA ECOSYSTEM

Nilsson, Palmér, and Naeve describe a Semantic Web metadata ecosystem in which metadata is conceptualized in terms other than the current environment of monolithic and authoritative records.[21] They describe a metadata ecology in which there is no canonical metadata record for a resource. Instead, the Web provides a distributed environment in which many independent statements about a resource coexist. Current work around the Dublin Core Abstract Model (DCAM), when wed with Semantic Web technologies such as Resource Description Framework (RDF), is intended to provide the means for enabling this new metadata ecosystem. The Semantic Web largely abandons the notion of the static, monolithic metadata description and instead defines a fluid environment of separate, but related resource descriptions more in keeping with the distributed nature of the Web and the communities building it.

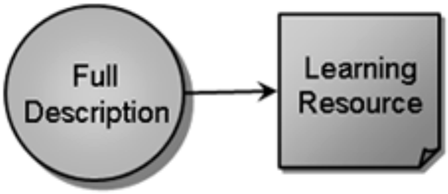
By static, monolithic resource descriptions we mean metadata records designed as a single, authoritative and presumed complete metadata record. According to Nilsson, this "view of meta-data is that it [sic] is something you produce once, often when you publish your document or resource, and which remains with the resource for its lifetime."[22] Traditionally, in the context of libraries, this unitary or monolithic metadata record was generated with great care by librarians trained in the principles and processes of cataloging. In general, the metadata records generated through these library processes are intended to be authoritative descriptions sufficient to serve the heterogeneous needs of library personnel and clientele.

In large part, the monolithic metadata records generated are intended to be canonical and shared. While it is true that there may be more than one such description of a given resource in existence, each of those descriptions is assumed complete and intended to stand alone. In such a context, metadata expressing the correlation of an educational resource to an achievement standard statement is just another part of the textual data making up the full or monolithic resource description. Such a record and its one-to-one relationship to the resource being described are illustrated in Figure 1.

In the library environment, notions of quality and the generation of quality metadata records of this monolithic sort have been an intrinsic function of the library community's efforts to produce canonical records according to established standards such as AACR and MARC for purposes of sharing descriptions and uniform processing and display. This push toward standard, quality description stemmed directly from the cooperative nature of librarians as a community of practice with shared values, interests, and fundamental goals. Bruce and Hillmann observe that this experience "in the library community reinforces two important points: quality is quantifiable and measurable and, and to be effective, enforcement of standards of quality must take place at the community level."[23]

In the early stages of the Web environment, general communities of practice emerged with specific descriptive needs including corporate/

FIGURE 1. Monolithic, Authoritative Description of the Learning Resource



enterprise, cultural heritage institutions (museums, libraries, and archives), education, geo-spatial, government, and recently social communities, to name a few. Studies by Lave and Wenger[24] and Wenger, Etienne, McDermott and Snyder[25] into the nature of communities of practice of this sort have demonstrated that the forms and cognitive processes behind purposeful actions vary substantially among practice communities. The inevitable result of this variance is that metadata necessary to meet the purposes of these communities varies as well–they each speak their own language or dialect and populate their value spaces with community jargon.

In general, the metadata models embraced by these emergent communities on the Web followed the monolithic forms of the pre-Web era. Examples include the IEEE LOM standard that defines a schema for describing learning objects and the Encoded Archival Description, to name two. Many non-standard monolithic schemas intended to serve a specific project or a federation of projects proliferated–some schemas migrating from purely local systems to the Web. As students of the Web know, this cacophony of community languages early on led to the Dublin Core Metadata Initiative (DCMI) with an initial goal of creating a pidgin language with a few general-use elements to support cross-domain resource discovery in this heterogeneous Web environment.[26]

Since work began on the Semantic Web in the late 1990s and the public attention drawn to it by the appearance in *Scientific American* of the Berners-Lee, Hendler and Lassila article,[27] theorists and practitioners of metadata have been reassessing the utility of monolithic forms of metadata in a networked environment. Driven largely by the very distributed nature of the Web and the potential of Resource Description Framework (RDF) with its capacity to frame meaningful relationships among descriptions of resources, the new metadata ecosystem as described by Nilsson, Palmér, and Naeve[28] and Downes[29] began to emerge.

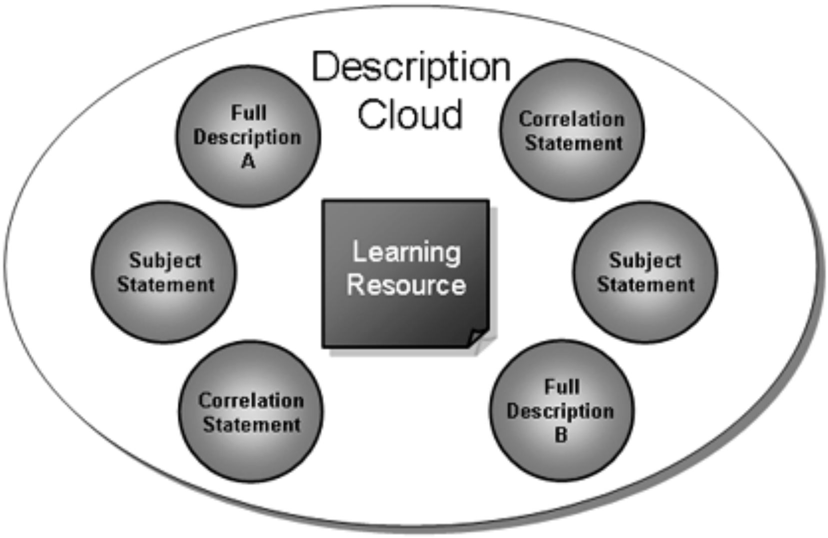
The general characteristics of metadata in the ecosystem are as follows:

- A single resource may have 1-*n* complete or fragmentary metadata descriptions reflecting differing ontological views as seen through the lenses of 1-*n* discourse and practice communities and individuals;
- These different points of view are manifest in 1-*n* schemas and schemes;
- The provenance of metadata statements is critical; and
- The fundamental metadata unit of concern is the single statement and not the record.

Figure 2 illustrates a number of these characteristics. In the figure we see a learning resource–perhaps a lesson plan addressing the Pythagorean Theorem–and a logical cloud of independent descriptions.

Some of the metadata statements or aggregations of statements in the form of full descriptions in Figure 2 may have been produced by highly reputable, readily recognized agencies while others may be peer standards correlations or other annotations where recognition of the individ-

FIGURE 2. Distributed Descriptions in the Logical Information Space



ual making the statement or his or her affiliation is limited. Thus, metadata statements with varying degrees of potential cognitive authority may be pulled together by applications and metadata aggregators from across the network using Semantic Web technologies and presented to a teacher as part of the logical information environment surrounding the resource. Having sufficient information regarding the correlator's identity and affiliation will be necessary to promote trust. Since each description in Figure 2–whether that description is a single statement or a set of statements–stands as a separate entity in the Semantic Web environment, relationships among them can be declared in RDF and used to aggregate them all into a composite–a logical record.

The logical information space illustrated in Figure 2 has an interesting corollary in work on metadata augmentation by Dushay, Hillmann and Phipps.[30] In their work with NSDL metadata harvested by means of OAI-PMH from a broad range of collection holders, Hillman, Dushay, and Phipps encountered the common problem faced by metadata aggregators in the Web environment of multiple descriptions of the same resource. Instead of looking to establish a canonical version of the metadata for the NSDL, Hillmann's team worked on a mechanism: (1) to combine all of the metadata from the various metadata instances describing a resource into a composite instance they call a "mudball"; and (2) to augment each metadata statement in the combined record with authority or provenance information. Dushay, Hillmann and Phipps posit that such a mudball can form the basis for framing what metadata would be presented to users by downstream applications in what they call "flavors."

In a limited sense, Dushay, Hillmann and Phipps' mudball fits what Downes calls a "resource profile":

> [T]his principle of resource profiles allows that the metadata for a given resource may be stored in different locations across the internet. That is, there is no single metadata file describing any given resource; metadata about the resource may be found in numerous online locations. A metadata profile is therefore constructed by aggregating the metadata available at these different locations in order to form a particular view of the resource. It follows that there may be different metadata profiles for a given resource, as different aggregators harvest different metadata from different locations, though one could define an ideal (and usually

fictional) "total" metadata profile composed of all possible metadata from all possible sources.[31]

Downes characterizes these forms of distributed metadata in terms of a tri-party system of metadata generation: (1) first-party metadata created by the resource author; (2) second-party metadata created by end users; and (3) third-part metadata created by a range of other agents (including librarians and other professionals). In a view of the metadata environment skewed toward the conventional framing of authoritative metadata, all metadata created by non-professionals are most frequently called "annotations"–some (mere) thing tacked onto the "real thing." We believe that such a perception is not shared by those guiding the development of Semantic Web principles and technologies. In the democratic nature of a network–where anyone can make statements about anything and "objectivity is defined by consensus"[32]–Downes' first-, second- and third-party metadata will coexist and interact as first class entities. Nilsson, Palmér, and Naeve speak of metadata on the Semantic Web in similar terms:

> In this scenario, meta-data for one resource need not be contained in a single RDF document. Translations might be administrated separately, and different categories of meta-data might be separated. Additional information might be added by others. Consensus building becomes a natural part of meta-data management, and meta-data can form part of the ongoing scientific discourse. The result is a global *meta-data eco-system*, a place where meta-data can flourish and cross-fertilize, where it can evolve and be reused in new and unanticipated contexts, and where everyone is allowed to participate.[33]

In Figure 3, we summarize and extend the thinking of Downes,[34] Nilsson, Palmér, and Naeve,[35] and Dushay, Hillmann and Phipps (2004).[36]

In the Figure, we see on the left-hand side our 1-*n* metadata statements and sets of statements as small circles distributed across the Web. These various statements are pulled together into what Dushay, Hillmann and Phipps call the mudball and Downes calls the resource profile. Once aggregated into the resource profile, metadata services may perform transformations and augmentations on the profile's component statements to improve quality and downstream utility. The resource profile can then be deployed by disseminators (metadata aggregators,

FIGURE 3. Dushay, Hillmann and Phipps' Mudball-to-Flavors–Resource Profile to Community Metadata Instances

end user applications, and other service providers) in the form of various "flavors" to discourse and practice communities based on community application profiles. One might imagine fairly simple applications in the future that enable disseminators to customize metadata flavors in response to personal profiles of users.

Given the circumstances of this emerging ecosystem, many quality indicators more suited to an environment of authoritative, monolithic metadata records become largely irrelevant in assessing metadata made available to an end user. Played out to its logical conclusion, the focus of what are identified in Table 1 as "intrinsic characteristics" remain applicable at the level of the individual metadata statement in the new ecosystem. However, the potential value of the "extrinsic (contextual) characteristics" shift from the composite metadata instance we traditionally call a record to the logical results of Semantic Web processes and other disseminators that gather, augment, and frame *logical* descriptions constructed from disparate metadata instances. Thus, assessing metadata quality within this ecosystem will be a task where our notions of *record completeness* or *conformance to community expectations* will be largely inapplicable directly to metadata instances where third-party statements were never intended to stand alone, but intended instead to exist in relation to other metadata instances.

It is to issues of metadata quality in the description of educational resources in this new metadata ecosystem that we now turn.

LEARNING RESOURCE CORRELATIONS TO ACHIEVEMENT STANDARDS

In the remaining sections of this article, we confine our discussion to metadata quality and utility issues as framed in terms of one attribute of resources that make them particularly useful for teachers and curriculum developers. Sutton posits that the statements that most distinguish the description of an educational resource from any other resource are statements about its audience characteristics, pedagogical aspects, and the learning objectives that are inherently or explicitly targeted by the resource.[37] Our discussion will focus on the last of these three characteristics. Thus, we are not concerned with the quality or utility of the general descriptive aspects of the resource such as title, author, publisher, and subjects, or the machine aspects of the resource including file types or content packaging.

We pick for discussion metadata statements about curriculum standards for two reasons: (1) curriculum standards have been generally mishandled in generating metadata describing learning resources; and (2) the current political climate in the U.S. around accountability in K-12 education as previously described demands that the metadata be more useful in supporting the teaching and learning enterprise.

The social and political thrust behind the national move toward accountability in K-12 education (and now beyond into higher education) has roots in our notion of standards-based systems of teaching and learning. Fundamental to this notion are "guiding questions":[38]

- *What knowledge and skills will students be learning?* This question is answered through the promulgation of achievement standards.
- *What experiences will be used to ensure that students learn?* The question is partially answered by identifying resources framing the teaching and learning experience.
- *What evidence will be gathered and used to ensure that students learn?* This question is answered through the assessment instruments used in the educational experience to assess learning.

In an effective education experience, there is a tight coupling between what is to be taught (question 1), how the students learn what is expected (question 2), and how that learning will be measured (question 3). The focus here is sharply on the second question identified by Gaddy, Dean and Kendall.

In daily practice, teachers use correlations of resources to standards for two separate, but related tasks:

1. *Information Search and Retrieval.* Consistently represented achievement standards can enhance information retrieval recall and precision by allowing a teacher to retrieve *all* and *only* resources addressing a specific achievement standard. Information about the learning objectives of a resource is frequently as important as its subject, and, on occasion, even more important. All teaching and learning activities include either explicit or implicit learning objectives that find expression in the promulgated achievement standards. Having the ability to target resources based on those standards can optimize limited teacher time spent searching. In addition, cross-mappings among similar standards in different jurisdictions enabled by unique identification and representation of standards makes it possible to expand searches to discover resources correlated to similar standards in other jurisdictions.
2. *Resource Compliance.* The political imperative for accountability in K-12 education means that teachers not only use the correlations for the general purposes of search and retrieval, but also to help them be accountable in their responsibilities to the children in their classrooms, the children's parents, the school, district and the nation. As a result, this task is permeated with issues of trust and levels of teacher confidence in the quality of a correlation. Research confirms that where the "consequences of use of information" or the "act or commitment based on information" are considered significant by the information seeker, they exercise heightened scrutiny in making information quality judgments.[39] An information seeker's trust is gained only when he or she sees the creator of the information as a "cognitive authority." Thus, cognitive authority is concerned with how people trust one another's opinions and is a conditional state of credibility bestowed by a trusting individual on someone or an agency. In essence, cognitive authority is rooted in identity and reputation and is concerned with "who knows what about what."[40] Thus, being able to identify the creator of the correlation and his or her affiliation will be necessary to building cognitive authority and trust.

These two tasks demand different levels of accuracy and precision in the correlation. A teacher might be quite satisfied for purposes of search and retrieval with correlations that manage to retrieve and to collocate resources *more* or *less* addressing an achievement standard. However, that teacher is likely to be dissatisfied with such results when looking for resources correlated with precision and authority in order to complete part of the fabric of his or her local curriculum. Thus, quality as "fitness for purpose" must be measured against both of these tasks.

LEARNING RESOURCE CORRELATIONS ON THE SEMANTIC WEB

As framed in our discussion of the emerging Semantic Web metadata ecosystem and illustrated in Figure 2, any number of third-party standard correlations will be generated describing different aspects of a specific resource. When provided with unique identifiers in the form of Uniform Resource Identifiers (URI), each of these disparate correlations has an independent existence on the Semantic Web and together form a logical information context as described by Downes in terms of the resource profile. [41] As we saw in Figure 3, any number of community-based application profiles can be derived from the aggregate resource profile.

In this context, the issues of quality–accuracy, physical accessibility, timeliness and provenance–adhere to the individual metadata statements from creation through possible aggregation in a resource profile and on into the context of various application profiles. Issues of utility–record completeness, intellectual accessibility, conformance to expectations, and consistency and cohesion–adhere to the varied results of the application profile disseminations in Figure 3 and are of less concern to the immediate discussion of metadata quality.

QUALITY STATEMENTS: ACCURACY AND THE VALUE STRING PROBLEM

One of the major problems to date in creating quality achievement standards metadata is the reliance on transcription of text strings in establishing the identity of a standard. With a few exceptions, the state jurisdictions in the U.S. have not put in place systematic mechanisms for identifying and referencing their standards documents. In those cases where the identity of the documents may be clear, the component statements that make up the taxonomic structures of those documents almost always lack unique identifiers. As a result, it is not uncommon to find value strings in metadata of the following sort:

> *Earth Sciences. 3. Weather can be observed, measured, and described. As a basis for understanding this concept: b. Students know that the weather changes from day to day but that trends in temperature or of rain (or snow) tend to be predictable during a season.*

Quite frequently, such a value string has been gleaned from an HTML page or a PDF document or been re-keyed with no small effort from a paper document. Many standards documents are difficult to interpret as a result of formatting and other decisions in their publication. These difficulties result in various interpretations across metadata statements of what a given standard actually is and how it should be represented as a value string. The vast majority of achievement standards metadata (including those in CEMARK records) substantially rely in this way on value strings to establish identity. This practice makes it impossible to reliably search and collocate shared resource descriptions based on their standards correlations. Even where effort has been made within a project to represent standards as value strings in a consistent manner, the data remains subject to error and is never interoperable with other projects where other interpretations hold.

Since the development of CEMARK in the early 1990s, the K-12 standards environment in the U.S. has grown increasingly complex. The assessments of what students should know and when they should know it are under continuous scrutiny at the national, state and district levels. A simple environment in which a state jurisdiction promulgates a set of enduring standards is unrealistic and unlikely. In the short period of the standards movement in the U.S., many states have produced more than one version of their state standards. In some instances, the versioning of these standards is evolutionary in the same sense that subject vocabularies in the bibliographic universe evolve as living languages develop. However, other versions of state standards are quite disjoint, representing radical shifts in thinking as the substance, methods and politics of K-12 education change. This complexity in the standards environment increases the need for providing clear identity to standards documents and their components standards statements.

In a few cases, states have been diligent in developing relatively sophisticated schemes for uniquely identifying achievement standards. For example, Ohio has its Standard Identifier Code (OSIC) by which the state identifies, with substantial clarity, taxon paths in its standards. For example, the OSIC identifier Y2003.CSC.S02.G11-12.BC.L12.I06 identifies the following taxon path (Ohio's textual labels have been added in italics for clarity):

» *Content Area:* Science

» *Standard:* 2. Heredity

» *Benchmark:* C. Explain how processes at the cellular level affect the functions and characteristics of an organism.

» *Indicator:* 6. Explain how developmental differentiation is regulated through the expression of different genes.

However, even given Ohio's efforts to establish clear identity for its standards, the identifiers are not guaranteed to be globally unique. In addition, the OSIC notation identifies the taxon path taken as a whole. There is no way within the Ohio scheme to identify individual statements within a single path–for example, identification of the example Ohio standard at the benchmark level as opposed to the indicator level. While this may not appear to be problematic in the Ohio example, it becomes problematic when the taxon paths are deeper and express a greater range of granularity from root to leaf. We will briefly explore this problem in a subsequent section of this article.

There are a number of proprietary systems such as Academic Benchmarks, EdGate, and Plato Learning that make encodings of standards available for use either in generating metadata in independent systems or for use within the proprietor's closed system of services and resources. Interoperability of achievement standards correlation metadata using these tools is not possible. Should one of these systems cease doing business as have MediaSeek and Align to Achieve, or the client simply wants to switch systems, data migration can be complex, time-consuming and costly. Of course, metadata generated by these systems suffers from the problems noted above–reliance on string values or non-global identifiers for the identification of standards and their component standards statements.

Where collection holders wish to share metadata describing learning resources more generally through harvesting mechanisms such as OAI-PMH, useful correlations must: (1) be based on a homogeneous, Web-friendly encoding of the content of the standards; and (2) use a global system of identification for both standards documents and each of the individual statements making up the documents' many taxon paths. The Achievement Standards Network (ASN),[42] developed in part for the National Science Digital Library, satisfies both of these requirements. In the ASN, over 450 current and historical state and nation U.S. standards documents have been "atomized" into nearly 800,000 individual statements–each uniquely identified with a Web de-referencable identifier in the form of a URI. De-referencing an ASN URI using an application or a Web browser returns an RDF/XML encoding of the text of all standard statements in the taxon path including brief metadata about the source document. A subset of information returned upon de-referencing an Ohio math URI is as follows:

> » Ohio Academic Content Standards K-12 Mathematics (2001) (*http://purl.org/ASN/resources/D100017A*)
>
> » Number, Number Sense and Operations Standard (*http://purl.org/ASN/resources/S1024934*)
>
> > » Computation and Estimation (*http://purl.org/ASN/resources/S100592F*)
> > Analyze and solve multi-step problems involving addition, subtraction, multiplication and division using an organized approach, and verify and interpret results with respect to the original problem. (*http://purl.org/ASN/resources/S1024B7C*)

The ASN defines its taxon path to include the statement identified by the ASN URI and all other statements in the upward path. Semantic and structural relationships among the atomic standards statements as well as the standards document are represented in RDF by means of their various resource URIs.

At a minimum, best practice in improving the intrinsic quality of correlation metadata requires the globally unique identification of the relevant standard statement and its associated taxon path through the assignment of the statement's URI. Whether the URI is de-referenced at the time the metadata is generated and the associated string values included in the metadata record or de-referenced in later operations on the metadata is an application issue beyond the scope of this article. However, de-referencing a statement URI and then assigning *only* the text of the taxon path in the metadata produces a substandard correlation. Even though such an assignment through the ASN would likely guarantee uniformity in text expression, such a metadata statement would not be operable in the Semantic Web context where rich relationships among resources depend on the URI.

QUALITY STATEMENTS: ACCURACY AND THE STRUCTURAL (GRANULARITY) PROBLEM

Kendall, in his paper titled "The Use of Metadata for the Identification and Retrieval of Resources for K-12 Education,"[43] identifies the problem of "fit" between the educational resource being correlated and a standard statement. Kendall frames his discussion in terms of the concept granularity of both the educational resource and the standard statement. In essence, Kendall asserts that the accuracy of the correlation is a function of the degree of granular fit between the two entities. To express this notion of strength of fit, we adopt the MARC 658 tag term "correlation factor."

Kendall's granularity problem is apparent in the complexity of the leaf statement in our preceding Ohio example: "Analyze and solve multi-step problems involving addition, subtraction, multiplication and division using an organized approach, and verify and interpret results with respect to the original problem." The statement can actually be factored out into sixteen distinct concepts as illustrated in Table 2.

An educational resource that addresses all sixteen concepts embedded in this Ohio standard would represent what we call a perfect "fit" between the resource and the standard–i.e., concept coverage between

TABLE. 2. Concept Breakdown of Ohio Canonical Statement

	Involving addition	**Involving subtraction**	**Involving multiplication**	**Involving division**
Analyze multi-step problems . . .	analyze addition	analyze subtraction	analyze multiplication	analyze division
Solve multi-step problems . . .	solve addition	solve subtraction	solve multiplication	solve division
Verify multi-step problems . . .	verify addition	verify subtraction	verify multiplication	verify division
Interpret multi-step problems . . .	interpret addition	interpret subtraction	interpret multiplication	interpret division

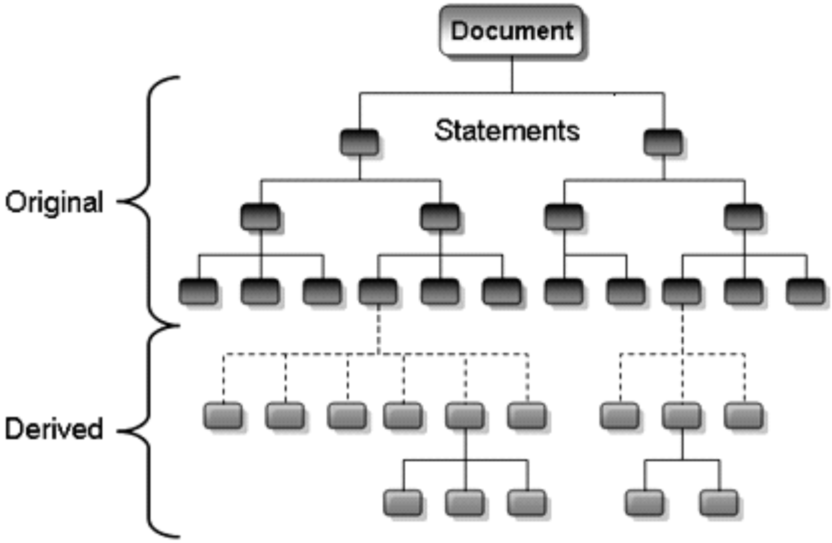
the two entities would be coextensive. In reality, however, it is highly unlikely that any single educational resource will address all sixteen concepts covered by the statement. For example, a resource may cover *analyzing* and *solving* but not *verifying* or *interpreting* or it may cover all of these cognitive tasks but only with regard to *addition* or *subtraction*. For such an educational resource, the fit is imperfect and any metadata assertion that the resource is useful in meeting the goals of the Ohio standard may be characterized as either inaccurate or very misleading. While such a fit might be acceptable for the *information search and retrieval task*, it would be quite unacceptable for the resource *compliance task*. We assert that many (if not most) educational resource correlations are of this sort–of imperfect fit. The existence of the MARC 658 Curriculum Objective subfield d for recording a "Correlation Factor" stands as a recognition of this general misalignment of fit between resource and standard.

However, to fulfill the *resource compliance task*, all correlations should push as close as possible to a perfect fit. There are two complementary mechanisms to assist in increasing metadata accuracy as it pertains to the problem of strength of fit: (1) by increasing the expressive power of the standards by providing for refinement of the canonical statements through the addition of more fine-grained standards statements; and (2) where a misalignment in granularity nevertheless exists, by increasing the expressive power of the correlation factor such that machines can unambiguously process metadata descriptions based on the strength of fit.

The first of these mechanisms–more granular expression of the standards–has been accommodated in the ASN architecture by providing a means to add more granular, derived statements. Such statements are assigned a status of derived because they represent derived refinements of the original standard statement by parties other than the standard's promulgating body. Just as the original statements in the ASN are identified by URI, so too are derived statements. Just as correlation metadata instances associated with a resource profile can be globally distributed in the Semantic Web ecosystem, so can these more granular standards statements be globally distributed, created by third parties, and aggregated into a standards document profile. Figure 4 illustrates such a profile.

While more granular expression of the standards will facilitate greater levels of accuracy in correlation assertions, it is highly unlikely that it will be possible to achieve a perfect fit in many instances. The second mechanism for improving the quality of correlations in terms of

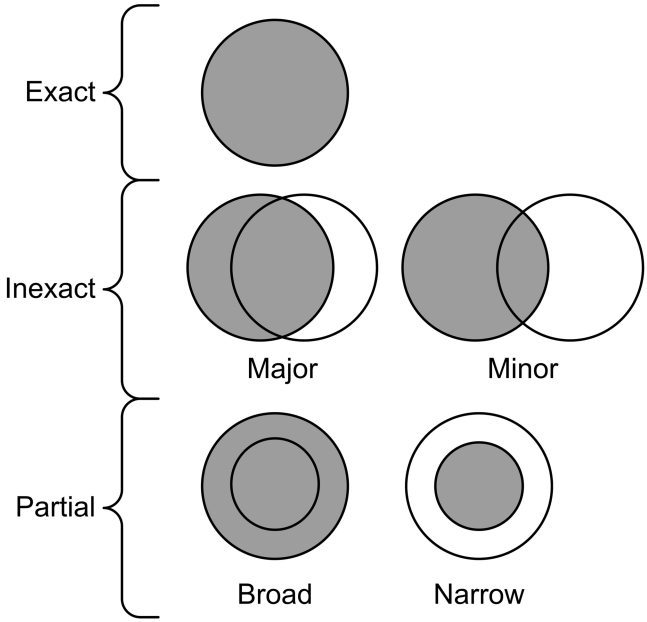
FIGURE 4. The Standards Document Profile Aggregating Original and Derived Standards Statements



strength of fit–machine amenable expression of a correlation factor–is intended to complement more granular standards statements by providing the means for applications to process the metadata based on machine assertions of the strength and nature of the fit. For example, an application might rank output of a search placing descriptions with a high level of fit at the top of displays or allow the user to eliminate from view resources below a certain threshold of fit.

Machine processing based on a correlation factor requires that different states of concept fit be unambiguously expressible in terms that support machine processing. To date, no standard, machine amenable correlation factor vocabulary has been developed. MARC 658 handles the expression of the factor by means of uncontrolled text strings–e.g., "slightly correlated," "moderately correlated" and "highly correlated"–terms that denote the strength but not the nature of the fit. National Science Foundation supported research is currently underway within the ASN project to develop a vocabulary addressing both the nature and strength of fit. At the time of this writing, the ASN project is using the inter-thesaurus mapping concepts modeled in Figure 5 as a base referent in developing an appropriate scheme with each resource with each state identified by a URI.

FIGURE 5. Machine Expression of Strength of Fit Using Adaptations of the SKOS Inter-Thesaurus Mapping Concepts



As noted, the MARC 658 tag has accommodated correlation factor data in bibliographic records since the creation of CEMARK in 1993. Neither the IEEE LOM nor the Dublin Core schemas support making correlation factor assertions. As a result, the binary nature of the standard correlation assertions using these schemas renders those assertions either inaccurate or very misleading.

THE CORRELATION ENTITY

We have tried to demonstrate so far that an assertion that a given resource correlates to a specific achievement standard is a complex statement demanding: (1) unambiguous identification of the standard statement, its associated taxon path and source document; (2) a sub-assertion regarding the nature and strength of fit between resource and standard concepts; and (3) a sub-assertion regarding the provenance of the correlation through which the user's perception of cognitive authority can be based.

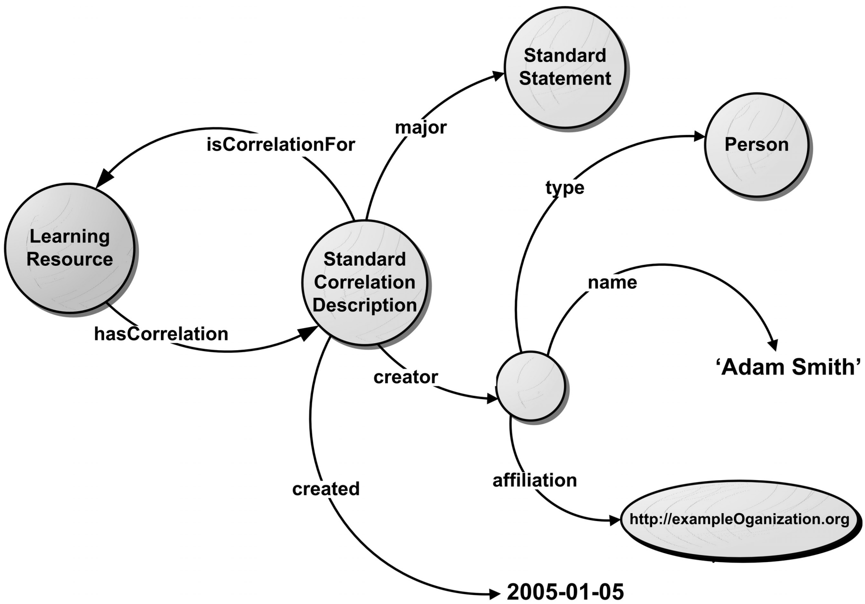
In closing, we note that in the context of the Semantic Web, these complex assertions can and should be modeled as separate resources that can accommodate and make this complex of information accessible to both machines and humans. In Figure 6, we graph such a standards correlation resource.

In the Figure we see each of the components of a quality correlation we have discussed. On the left we see the resource being described–for example, an individual lesson plan. The standard correlation resource in the center of the Figure is the subject of both a correlation assertion that denotes the correlation factor ("major"), and an authorship statement upon which cognitive authority can be based.

CONCLUSION

Correlation of educational resources to achievement standards is now a fundamental requirement in metadata generation for the K-12

FIGURE 6. The Standards Correlation Resource on the Semantic Web



community of practice. In this article, we have explored the nature of such correlations in terms of metadata quality for descriptions that will serve the dual needs of K-12 teachers–*search and retrieval* and *resource compliance*. In the course of the discussion, we reach the conclusion that the emerging Semantic Web metadata ecosystem asks us to reevaluate our notions of metadata quality in the new environment and to frame our measures in terms of intrinsic attributes informing a more objective notion of *quality* and extrinsic attributes framing a more general, subjective notion of *utility*. Within this dual framework, we determined that the generation of quality correlation metadata requires us to: (1) abandon our traditional reliance in the bibliographic community on value strings as the primary means of identifying achievement standards in metadata; (2) provide appropriate means to identify the provenance of all correlations assertions for purposes of cognitive authority; and (3) solve the problem of misalignment in terms of granularity between resource and declared objectives by: (a) generating more granular achievement standards statements; and (b) developing and applying machine- and human-readable correlation factor information.

Received: April, 2007
Revised: June, 2007
Accepted: September, 2007

REFERENCES

1. Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform* (Washington, DC: U.S. Government Printing Office, 1983).

2. For background on Curriculum-enhanced MARC, see Lorraine Knight, "The 658 Story" (paper, The National Curriculum Enhanced MARC Conference, Bowling Green, Ohio, 1993), http://www.nmm.net/market_resources_658_story.shtml; Catherine Murphy. *Curriculum-Enhanced MARC (CEMARK): A New Cataloging Format for School Librarians*, ERIC document no. ED399952 (Washington, D.C.: Educational Resource Information Center, 1995); and Tom Adamich, "Curriculum-enhanced MARC," *Knowledge Quest on the Web*, 35 (September/October 2006): 65-68, http://www.ala.org/ala/aasl/aaslpubsandjournals/kqweb/kqarchives/volume35/KQW35_1Adamich.pdf.

3. William E. Moen, Erlin 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," http://www.unt.edu/wmoen/publications/GILSMDContentAnalysis.htm.

4. Thomas R. Bruce and Diane I. Hillmann, "The Continuum of Metadata Quality: Defining, Expressing, Exploiting," in *Metadata in Practice*, eds. Diane I. Hillmann and Elaine L. Westbrooks (Chicago: American Library Association, 2004), 238-256.

5. Roy Tennant, "Specifications for Metadata Processing Tools," http://www.cdlib.org/inside/projects/harvesting/metadata_tools.htm.

6. Jane Barton, Sarah Currier and Jessie M. N. Hey, "Building Quality Assurance into Metadata Creation: An Analysis Based on the Learning Objects and E-prints Communities of Practice," in *Proceedings of the International DCMI Metadata Conference and Workshop*, 2003, 39-48, http://www.siderean.com/dc2003/201_paper60.pdf.

7. Besiki Stvilia, Les Gasser, Michael B. Twidale, Sarah L. Shreeves, and Tim W. Cole, "Metadata Quality for Federated Collections," in *Proceedings of the International Conference on Information Quality*, 2004, 111-125.

8. Paul Johanis, "Assessing the Quality of Metadata: The Next Challenge" (presentation, Working Session on METIS, Luxembourg, 2002), http://www.unece.org/stats/documents/2002/03/metis/19.add.2.e.pdf.

9. Marieke Guy, Andy Powell and Michael Day, "Improving the Quality of Metadata in Eprint Archives," *Ariadne* 38 (2004), http://www.ariadne.ac.uk/issue38/guy/.

10. Stvilia et al., "Metadata Quality."

11. Les Gasser and Besiki Stvilia, "A New Framework for Information Quality. Technical Report ISRN UIUCLIS-2001/1 + AMAS" (technical report, University of Illinois at Urbana Champaign, 2001).

12. Johanis, "Assessing the Quality."

13. Bruce and Hillmann, "Continuum of Metadata," 242-49.

14. Moen, Stewart, and McClure, "Role of Content Analysis."

15. Guy, Powell and Day, "Improving the Quality."

16. Stvilia et al., "Metadata Quality," 114.

17. Bruce and Hillmann, "Continuum of Metadata," 244-45.

18. Mikael Nilsson, Matthias Palmér, and Ambjörn Naeve, "Semantic Web Metadata for E-Learning–Some Architectural Guidelines," in *Proceedings of the 11th World Wide Web Conference*, 2002, http://kmr.nada.kth.se/papers/SemanticWeb/p744-nilsson.pdf.

19. Bruce and Hillmann, "Continuum of Metadata," 244.

20. Stvilia et al., "Metadata Quality."

21. Nilsson, Palmér, and Naeve, "Semantic Web Meta-data."

22. Ibid.

23. Bruce and Hillmann, "Continuum of Metadata," 240.

24. Jean Lave and Etienne Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge: Cambridge University Press, 1991).

25. Etienne Wenger, R. L. McDermott and W. Snyder, *Cultivating Communities of Practice* (Cambridge, MA: Harvard Business School Press, 2002).

26. Thomas Baker, "A Grammar of Dublin Core," *D-Lib Magazine* 6, no. 10 (2000), http://www.dlib.org/dlib/october00/baker/10baker.html.

27. Tim Berners-Lee, James Hendler and Ora Lassila, "The Semantic Web," *Scientific American* (May 2001): 34-43, http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21.

28. Nilsson, Palmér, and Naeve, "Semantic Web Meta-data."

29. Stephen Downs, "Resource Profiles. Version 3," (2003), http://www.downes.ca/files/resource_profiles.htm.

30. Naomi Dushay, Diane I. Hillmann, and Jon Phipps, "Improving metadata quality: Augmentation and recombination," in *Proceedings of the International DCMI Metadata Conference and Workshop*, 2004, http://www.slais.ubc.ca/PEOPLE/faculty/tennis-p/dcpapers2004/Paper_21.pdf.

31. Downs, "Resource Profiles."

32. Nilsson, Palmér, and Naeve, "Semantic Web Meta-data."

33. Ibid.

34. Downs, "Resource Profiles."

35. Nilsson, Palmér, and Naeve, "Semantic Web Meta-data."

36. Dushay, Hillmann, and Phipps, "Improving Metadata Quality."

37. Stuart A. Sutton, "Principled Design of Metadata-generation Tools for Educational Resources," in *Developing Digital Libraries for K-12 Education*, ed. by Marcia A. Mardis (Syracuse, NY: ERIC Clearinghouse on Information and Technology, 2003): 45-63.

38. Barbara B. Gaddy, Ceri B. Dean, and John S. Kendall, "Keeping the Focus on Learning," *Noteworthy Perspectives* (Aurora, CO: Mid-Continent Research for Education and Learning, 2002), http://www.mcrel.org/PDF/Noteworthy/5022IR_NW_Focus.pdf.

39. Soo Young Rieh, "Judgment of Information Quality and Cognitive Authority in the Web," *Journal of the American Society for Information Science and Technology* 53, no 2 (2002): 145-161.

40. Patrick Wilson, *Second-hand Knowledge: An Inquiry into Cognitive Authority* (Westport, CT: Greenwood Press, 1983).

41. Downs, "Resource Profiles."

42. GEM Exchange, "NSDL: Achievement Standards Network," http://www.thegateway.org/asn.

43. John S. Kendall, "The Use of Metadata for the Identification and Retrieval of Resources for K-12 Education," in *Proceedings of the International DCMI Metadata Conference and Workshop*, 2003: 109-117, http://www.siderean.com/dc2003/304_Paper68.pdf.