The Dublin Core and Metadata for Educational Resources

Stuart A. Sutton, The Information School, University of Washington
Seattle, WA, USA
sasutton@u.washington.edu

Jon Mason, EdNA, *education.au limited*
Dulwich, Australia
jmason@educationau.edu.au

Abstract

This paper focuses on metadata standards for networked information discovery and retrieval of educational resources. Such standards, if they achieve a high degree of interoperability, will support both extant and emergent e-learning architectures. This paper takes an historical perspective on the formative work of the Dublin Core Metadata Initiative Education Working Group and explores issues remaining for future resolution. The paper also examines the relationship between Dublin Core Metadata Initiative and the IEEE Learning Technologies Standards Committee Learning Object Metadata Working Group.

Keywords: Educational Metadata, Networked Information Discovery and Retrieval, Dublin Core Element Set, IEEE LTSC LOM, Dublin Core Education Working Group

1. Introduction

In this paper, we explore along a number of dimensions both the underlying principles of the Dublin Core Metadata Element Set (DCMES)[1] and how those principles are playing out in the description of educational resources for networked information discovery and retrieval. We first set out an array of fundamental principles that distinguish the work of the Dublin Core Metadata Initiative (DCMI) from that of others as a framework for defining the intentions behind the work of the DCMI Education Working Group (DC-Ed).[2] Next, we discuss the classes of metadata statements necessary to support networked information discovery and retrieval in the educational domain. We then discuss the DCMI Usage Board recommendations stemming from the DC-Ed initial set of proposals as well as the nature of those proposals as an application profile. We conclude by defining a number of possible next-steps for the Working Group.

It is important that we draw a clear distinction here at the beginning between the objectives of projects purposed in networked information discovery and retrieval and projects whose purposes rest in the more general management and deployment of learning objects—e.g., the IMS Global Learning Consortium (IMS) [3] and ARIADNE. [4] This distinction is important for a number of reasons. First, it is intended to emphasize the broader scope of standards development of the latter-mentioned projects, a scope that attempts to be holistic in specifying the technical components and requirements of systems that support online education. Second, to defuse any misconception that the standards are competitive solutions to a single problem. And third, to highlight the special focus of networked information discovery and retrieval, a focus that brings with it a perspective that can inform the wider agenda involving the design of online instruction.

The paper focuses heavily on the work of both existing DCMI-based metadata projects in the educational domain, on the work of DC-Ed and the relationship between DCMES and the IEEE LTSC LOM Working Group. In particular, two projects—Education Network Australia (EdNA) [5] and Gateway to Educational Materials (GEM) [6]—are used throughout as exemplars of various issues.

2. Guiding Principles

The metadata projects generally described in this paper make up a very specific community of practice—a community concerned with networked information discovery and retrieval of educational resources, or 'objects', that have developed their data models relying on a series of principles defined by the DCMI. This is not to imply that the other models (e.g., IEEE P1484.12 (LOM) [7]) do not contain elements whose semantics map more or less successfully to the DCMES. However, to assert that a project is "DCMI-based" within the meaning of this paper requires adherence to both the DCMI namespace and the guiding principles that define DCMI's conceptual framework. Throughout this paper, we use the term "DCMI namespace" to denote one or more schema defined and managed by DCMI under its registration authority.

The DCMI principles of import to this discussion include: (1) the extensibility of the DCMES; (2) mechanisms for the constrained qualification of elements and their accompanying values based on DCMI principles; and (3) modularity—the notion of multiple metadata instances relating to a single resource. Since these principles frame the DC-Ed proposal and the DCMI recommendations, each will be briefly chronicled here with citations to fuller explanations.

The primary goal of the DCMI framework is to create a broad, interdisciplinary, international consensus around a core set of elements that are useful in describing a wide array of resources across the Internet, and useful to a diversity of communities or domains of interest for the purpose of networked information discovery and retrieval. While it is recognized that there are many needs for metadata beyond the goal of networked information discovery and retrieval, those needs fall outside the DCMI agenda and do not directly inform its work.

Principle: Element and Value Qualifiers. From its beginnings in 1995, the DCMI founders recognized that while the fifteen elements and their core semantics would be useful for coarse-grained, cross-domain networked information discovery and retrieval, the element set would be insufficient to serve more precise information retrieval needs. As a result, the DCMES can be enriched through use of qualifying mechanisms known as the "Canberra Qualifiers." [8] These qualifier mechanisms make it possible to refine an element in meaningful ways and to constrain values assigned to it.

Principle: Extensibility. In addition to being able to qualify elements and element values, the principle of extensibility permits creation of additional elements (and element qualifiers) where necessary to meet particular metadata needs of local or domain specific applications. For example, nearly every education-specific metadata set rooted in the principles of the DCMI has created an element of one sort or another to express various characteristics of the "audience" for which the resource being described is either intended or otherwise useful. The DCMI principle of extensibility permits creation of such a "local" element and accompanying qualifiers.

Principle: Interoperability. However, it goes without saying that such local elements raise issues of cross-domain or even intra-domain interoperability. For example, in the ongoing process of achieving a national education metadata standard for education in Australia, a number of state education authorities articulated schemas during 1998-99 that posed challenges to the harmonization required for a national standard. Local identity and "ownership" of resources to be made available in the aggregated pooling on EdNA Online [9] became an issue partly as a result of the legacy that each state education system has developed different approaches to standards in curriculum and assessment.

Three levels of semantic interoperability are identifiable in the DCMES framework:

- *Level 1—Cross-Domain:* Entities that are judged to be widely useful across disciplines and communities of practice; i.e., they are *core* entities (DCMI namespace);
- *Level 2—Domain-Specific:* Entities that are judged to be useful within a domain, but perhaps not across domains (DCMI namespace); and
- *Level 3—Local:* Entities that are useful for local applications or in a constrained federation of applications, but, perhaps, are not widely used even in a given domain (varying namespaces).

Obviously, the three semantic interoperability levels represent a descending order of cross-domain interoperability. With the proliferation of elements at Levels 2 and 3 and the translation of those elements into multiple languages, the potential meaning in element names will be assuredly lost absent some public, machine accessible way for managing the evolving DCMI namespace. The DCMI registry under development promises such management.

Under principles for managing the DCMI namespace currently emerging, an element or qualifier that conforms to the constraints of the DCMI principles outlined earlier is considered conforming and may be "assigned a global interoperability token within the DCMI namespace." [10] In addition, such an element or qualifier is assigned a status within the namespace of either "Cross-Domain" or "Domain-Specific."

Providing access to elements and their accompanying qualifiers and semantics in the DCMI namespace is one function of the emerging DCMI registry.

Principle: Metadata Modularity. It was understood in the DCMI community early on that no single metadata schema would satisfy all needs and that the information environment surrounding a particular resource might be filled with disparate, complementary metadata packages. This principle of modularity was first fully articulated through the Warwick Framework [11]. One of the basic purposes of the Warwick Framework has been recently restated in terms of the notion of "application profiles". Application profiles are defined as "schemas which consist of data elements drawn from one or more namespaces, combined together by implementers, and optimised for a particular local application." [12] As we shall see later, the DCMI Education Working Group proposal is a form of application profile.

Thus, we see through applying the principles and mechanisms of extensibility and metadata modularity a possible means for tailoring metadata to meet the needs of discourse and practice communities such as education. In the DC-Ed proposal, and, to a certain degree in the DCMI recommendations, we see these mechanisms at work. It is to a discussion of these proposals and recommendations that we now turn.

3. DC-Ed

In August, 1999, the DCMI Advisory Committee formed the Education Working Group (DC-Ed) with the charge "to discuss and develop a proposal for the use of Dublin Core metadata in the description of educational resources." Creation of the DC-Ed Working Group marked a sharp departure from DCMI practice in that its charge involved the exploration of possibly new domain-specific DCMI elements as well as qualifiers.

Since the mid-1990s, various metadata projects around the world developed element sets for describing educational resources. A number of the projects were based on DCMI principles. However, these DCMI-based projects pursued different paths in their use of the extensibility and qualifier principles. For example, the Gateway to Educational Materials (GEM) project in the United States relied heavily on the use of element qualification while holding the number of new, domain specific elements to a minimum.[13] In contrast, Education Network Australia (EdNA) did not initially create element qualifiers but reached expressive ends similar to GEM through the use of more elements. In developing its schema, EdNA took the view that unqualified elements would likely provide a more practical basis toward interoperability for its stakeholders, and at the time, there was little consensus in the DCMI regarding implementation of qualifiers for the DCMES.

EdNA's decision to pursue its own domain specific elements, however, took place with the proviso that revisions or updates would only take place in harmony with either formal DCMI revisions or with Australian whole-of-government initiatives.[14] Thus, while rooted in the principles of the DCMI, GEM and EdNA's different developmental paths and modeling choices meant little interoperability beyond that provided by unqualified DCMES.

In brief, the original goal of DC-Ed was the development of a common set of DCMI recommended elements and qualifiers for use with educational materials that promote interoperability among projects such as EdNA Online and GEM. Through the DCMI framework for interoperability described above, this common set of domain-based DCMI elements and qualifiers make possible the following: (1) full Level 1 interoperability through the DCMES; (2) shared domain-specific elements and qualifiers for Level 2 interoperability for educational resources; and (3) the freedom to develop local elements and qualifiers as needed at interoperability Level 3.

3.1 Common Educational Metadata Needs

After its formation in August 1999, the Working Group performed two separate analyses of existing metadata projects and standards efforts in the educational domain. While many of the projects analyzed were DCMI-based, several initiatives (e.g., ARIADNE and IMS) were more directly rooted in the emerging IEEE LTSC Learning Object Metadata (LOM) standard.

One of the goals of these analyses was to ferret out domain-specific elements with more-or-less common semantics. For this work, an admittedly crude metric for "commonality" was used. In order for an attribute (however named) to be considered "common," an element expressing that attribute had to show up in two or more project element lists. Examination of the entries in this listing of commonly held attributes revealed four general classes of attributes useful for networked information discovery and retrieval that are not expressible through the fifteen DCMES. These common classes of domain-specific attributes are summarized as follows: [15]

- *Audience.* A broad class containing various attributes describing characteristics of the target audience (or audiences) for a resource ranging from academic level (e.g., "5th grade") through physical, emotional, social, and intellectual characteristic.
- *Duration.* Information denoting the normal "use" time for the resource as opposed to any physical run time.
- *Learning Processes/Characteristics*. A broad class denoting various activities and methods used by an instructor or trainer including those stemming from various teaching/learning theories, student groupings, and assessment methods, etc.
- *Competencies.* A class of educational goals or objectives that may be either particular (i.e., local) to the resource being described, or they may be formally promulgated national, international, or organizational content/process standards for which a resource was either intended or for which it is deemed useful by the creator, publisher, or third party.

Based on its research and discussions, DC-Ed issued an initial proposal for element and element qualifiers that laid the foundation for addressing in greater depth in the future a number of the identified classes of metadata noted above. [16]

3.2 Proposals and Recommendations

In the DCMI, the Usage Board is charged with reviewing select proposed modifications to the DCMI namespace including review and action on proposed new elements and qualifiers. The results of its deliberations are either a DCMI recommendation or a refusal to recommend with an accompanying explanation.

DCMI recommendations with regard to the addition of new, or the modification of existing elements and qualifiers become part of the DCMI namespace. On occasion, the Usage Board may issue a recommendation rendering an element or qualifier in the DCMI namespace obsolete. In instances other than those of obsolescence, the Usage Board assigns a recommendation status of either "Cross-Domain" or "Domain-Specific" reflecting the level of interoperability as noted earlier.

In the following sections, we discuss the DCMI recommendations stemming from the initial DC-Ed proposal. [17]

3.2.1 Audience Proposal

One of the most common needs across the projects analyzed by the Working Group was the capacity to make statements about the various aspects of the intended users of the educational resource being described. Frequently, creators and publishers of resources explicitly state the type or class of user for whom the resource is intended or useful. As a result, DC-Ed proposed a new element for the DCMI namespace with the name "Audience." The Usage Board issued a recommendation with a "Domain-Specific" status:

DCMI Element Recommendation
Recommendation Status: Domain-Specific
Name: audience
Label: Audience
Definition: A class of entity for whom the resource is intended or useful.
Comment: A class of entity may be determined by the creator or the publisher or by a third party.

As a class, audience may be subdivided into two general subclasses that include in the first instance those persons, organizations, and other forms of entities that administer or mediate access to a resource by a second subclass of audience–the "end" users for whose benefit the resource being described was designed.

Based on the DC-Ed proposal, the Usage Board issued a DCMI Recommendation for an "Audience" element qualifier named "Mediator" to hold expressions of the first general subclass. It was assigned "Domain-Specific" status:

DCMI Audience Element Qualifier Recommendation
Recommendation Status: Domain-Specific
Name: mediator
Label: Mediator
Definition: A class of entity that mediates access to the resource and for whom the resource is intended or useful.

3.2.2 Competency Proposal

It was unanimously agreed among members of the Working Group that the capacity to associate the educational resource being described with organizational, professional, province/state, national, and international content and process standards is an important function for networked information discovery and retrieval. The Working Group identified two different implementations achieving the goal: (1) definition of a new education-specific element named "Standard", and (2) use of the existing DCMES "Relation" element with a new element qualifier named "Conforms To." The Usage Board issued a recommendation for the additional qualification of the existing "Relation" element. The Board did not recommend a "Standard" element for the DCMI namespace since the goals of the proposal as framed could be as easily met using the "Relation" qualifier. "ConformsTo" was assigned "Domain-Specific" status:

DCMI Relation Element Qualifier Recommendation
Recommendation Status: Domain-Specific
Name: conformsTo
Label: ConformsTo
Definition: A reference to an established standard to which the resource conforms.

4. The DC-Ed Application Profile

In December of 2000, DCMI and the IEEE Learning Technologies Standards Committee (LTSC) LOM Working Group released the text of a memorandum of understanding (MOU) defining a cooperative relationship between the two bodies. [18] A number of projects including ARIADNE, EdNA, GEM, and IMS concurred in the memorandum.

The intention of the MOU was for the parties to work together to minimize "barriers to the creation, interchange, and use of metadata." In the words of the MOU: "Based on this philosophy, we agree to support modular, extensible, structured metadata."

To understand the basis for the MOU, it is necessary to understand the functional relationship between the two schemas; and, to do so, it is necessary to understand several of the functions of metadata in the education and training domains.

Risking over-simplification, we posit two fuzzy classes of resources in the digital information environment that supports education and training. The first class of resources are those that were educationally-purposed from inception—a class or resources many call learning objects. In addition to discovery (since you cannot use what you cannot find),

metadata requirements for the learning object class are with the automatic management and use of those objects within learning management systems (LMS). The standards work of the IEEE LTSC (and its affiliated projects) focuses on the broad array of issues for this class including content models and object-LMS interactions.

The second class of resources that can be useful in the processes of teaching and learning are those resources not educationally-purposed at inception but may have educational value when repurposed for that context. For example, weather data collected from a specific instrument for use in weather prediction services may be utilised within a weather simulation tool for K-12 instruction. Vast numbers of such resources exist in repositories across the Internet. The metadata work of the DCMI (and its affiliated projects) focuses on the discovery and retrieval of such resources across disparate domains.

It is growing increasingly clear that collection holders of educationally-purposed learning objects want to expose those objects to information systems outside the context of Learning Management Systems (LMS) and repository holders of non-educationally-purposed resources want to expose their resources to LMS. [19] Thus, we see the potentially synergistic relationship between DCMI and the IEEE LTSC LOM Working Group foreshadowed in the MOU.

In August 2001, the signatories and concurring parties to the MOU met in Ottawa to chart an agenda for its operationalization. The problems to be addressed in that work include, but are not limited to the following:

- Dispelling the general confusion in the education marketplace about the various functions of metadata—particularly confusion regarding the functions of DCMES in cross-domain resource discovery and the functions of the IEEE LTSC LOM in the management and deployment of learning objects;
- Dispelling misplaced perceptions that DCMES and IEEE LTSC LOM are competitive solutions to the same problem—i.e., assumptions that they address the same functions;
- Dispelling confusion in the marketplace over the issues of combining complementary but independent metadata element sets in application profiles in order to meet the disparate needs of individual and federated projects; and
- Providing solutions to issues of technical interoperability between the DCMES and IEEE LTSC LOM schemas through: (a) development of an architecture for application profiles integrating some or all aspects of the two schemas, (b) clear syntactic bindings for metadata instances using those application profiles, and (c) collaboration on common metadata registry issues.

In the spirit of the MOU and in addition to its proposal for additional domain-specific elements and qualifiers within the DCMI namespace, the Education Working Group also endorsed three elements drawn from the IEEE LTSC LOM namespace: (1) Typical Learning Time, (2) Interactivity Type, and (3) Interactivity Level. One of these IEEE LTSC LOM elements—Typical Learning Time—fulfills the need identified by the Working Group in its proposal for a "Duration" element as a means of making statements about the useful use time of the resource in a teaching and learning context. As a result of the use of elements from multiple namespaces (DCMI and IEEE LTSC LOM), the DC-Ed proposal is a nascent application profile as defined by Heery and Patel. [20]

4.1 Future Developmental Paths

Future work to meet the DC education community's needs will most likely evolve along three separate but inextricably related paths.

First, where it is possible to make the needed metadata statements through terms in established standards (e.g., IEEE LTSC LOM) those needs can be fulfilled by referencing those terms in their associated namespaces and defining the nature of their use through the DC-Education Application Profile.

Second, those needs might be met through the Application Profile by referencing elements and qualifiers in established, well-maintained namespaces other than those of the standards promulgating bodies (e.g., GEM and EdNA).

Third, where there is broad consensus among the DC education community that an element or qualifier is needed that does not already exist in a recognized schema, or that needs to be managed by the DCMI community, then the processes developed for the DC-Ed initial proposal may be used to devise new domain-specific elements and qualifiers for the DCMI namespace.

4.2 Future Development Issues

The DC-Ed work leading to the initial proposals and recommendations was intended to lay the foundation for future work of the Working Group. Two of the classes of metadata statements originally identified present significant challenges—"Audience" and "Learning Processes/Characteristics."

A number of significant facets of the new "Audience" element that emerged through Working Group discussions, examination of existing practice, and from authoritative sources such as the *Thesaurus of ERIC Descriptors* need greater scrutiny. They include:

- *Intelligence/Ability.* Used for an audience with specific intellectual characteristics; e.g., gifted students.
- *Physical/Emotional.* Used to denote physical, learning, and developmental disabilities of the target audience; e.g., visually impaired students or students with attention deficit disorder.
- *Cultural/Linguistic Group.* Groups for whom the resource is intended; e.g., Asian, African, and Mexican Americans.
- *Education/Training Level.* A general statement describing the education or training sector. Alternatively, a more specific statement of the location of the audience in terms of its progression through an education or training sector.
- *Gender.* The gender of the audience for whom the resource was designed.
- *Linguistic Ability.* The native language of the target audience (which may be different from the encoding language of the resource).
- *Knowledge Prerequisite.* Competencies a student is assumed to have achieved as a prerequisite for successful use of the resource.

Exploration of these facets is ongoing in DC-Ed and will likely shape subsequent proposals for qualification of the "Audience" element. For example, there is already a DC-Ed consensus that the "Education/Training Level" class is needed as an "Audience" element qualifier. It is likely that work going on within other relevant fora, such as the W3C and IMS, will also inform DC-Ed requirements particularly where accessibility of learning resources need to be well-described.

In discussions since the Usage Board action on the DC-Ed proposal, it has become clear that the new "Relation" qualifier, "conformsTo" only partially meets the goal of competency descriptions. While it serves well as a reference to an external content standard, it is inappropriate for statements about intended competencies that are local to the resource being described. Additional discussion and Working Group action is likely in this area.

The thorniest of the classes yet to be tackled by DC-Ed is the Learning Processes and Characteristics class. While admittedly difficult, it is on the Working Group agenda. Again, it is likely that work currently underway within the IMS Learning Design Working Group may also indicate the way ahead.

5. Conclusion

Some further reflection upon contextual issues should underscore the developmental character of the DC-Ed Working Group's initial proposals. In terms of process, there is other commentary that has relevance where a wide scoping of the issues concerning the *design* of teaching and learning is involved. While the charter of DC-Ed is not explicitly concerned with matters such as the design of teaching and learning, its focus on networked information discovery and retrieval for educational resources is not without context. Networked information discovery and retrieval is probably recognized by many DC-Ed practitioners as a key activity that supports, in a broad sense, both e-learning and the emerging education object economy.

As the so-called "knowledge age" evolves, classifications and standards will proliferate. They will be developed at the interface where technical capabilities and requirements are considered in tandem with the socio-political requirements of a particular community of practice (in this case, education). This is not to say that each community of practice will operate only according to its own rules—to the contrary. Interoperable metadata demands cooperation at the intersections of multiple communities of practice.

In this paper, we have chronicled the evolution of educational metadata within the context of the DCMI. We are at the point in this history where disparate metadata projects joined at the margins through DCMI principles, are seeking to coalesce as a coherent community of practice in order to promote global interoperability for networked information discovery and retrieval in the domain of educational objects. While the original mandates of the DCMI-based projects were largely parochial in nature—serving specific nations or individual sectors within nations, the DCMI goal of international consensus demands a high level of cooperation among these projects. The core principles and processes of the DCMI support such cooperation. The creation of the DC-Ed Working Group was the first major step in achieving a level of consensus that will bind this emerging community.

Notes and References:

[1] http://www.dublincore.org/documents/ dces/
[2] http://www.dublincore.org/groups/ education/
[3] http://www.imsproject.org/
[4] http://ariadne.unil.ch/
[5] http://www.edna.edu.au/
[6] http://www.TheGateway.org
[7] http://ltsc.ieee.org/wg12/
[9] http://www.edna.edu.au/
[10] T. Baker. (January, 2000) "A Multilingual Registry for Dublin Core Elements and Qualifiers," *Zeitschrift fuer Bibliothekswesen und Bibliographie* Available: http://www.mailbase.ac.uk/lists/dc-registry/1999-12/0000.html
[11] L. Dempsey & S. Weibel. (1996). The Warwick Metadata Workshop: A framework for the deployment of resource description, *D-Lib Magazine* (July/August) Available: http://www.dlib.org/dlib/july96/07weibel.html; see also, C. Lagoze, C. Lynch, & R. Daniel Jr. (1996). *The*

Warwick Framework: A container architecture for aggregating sets of metadata. TR96-1593 Available: http://www.nlc-bnc.ca/ifla/documents/libraries/cataloging/metadata/tr961593.pdf
[12] R. Heery & P. Manjula. (2000). Application profiles: Mixing and matching metadata schemas, *ARIADNE*. Available: http://www.ariadne.ac.uk/issue25/app-profiles/
[13]http://geminfo.org; http://www.thegateway.org
[14] The AGLS (Australian Government Locator Service) is a DCMI-based schema with a small set of extensions/qualifiers. It is mandated throughout all Australian Government departments and agencies.
[14] Education Working Group. (2000). *Report of Deliberations.* S. A. Sutton & J. Mason, Eds. Available: http://www.ischool.washington.edu/sasutton/dc-ed/Dc-ac/DC-Education_Report.html
[16] http://www.dublincore.org/documents/2000/10/05/education-namespace/
[17] http://www.dublincore.org/documents/#recommendations
[18] http://dublincore.org/documents/2000/12/06/dcmi-ieee-mou/
[19] LMS are sometimes referred to as Managed Learning Environments (MLE)
[20] R. Heery & P. Manjula. (2000). Application profiles: Mixing and matching metadata schemas, *ARIADNE*. Available: http://www.ariadne.ac.uk/issue25/app-profiles/