DC Metadata is Alive and Well
(and has Influenced a New Standard for Education)

Liddy Nevile
Australia
liddy@sunriseresearch.org

Abstract
This paper describes the emerging ISO/IEC 19788 Metadata for Learning Resources standard, where DCMI specifications have been adopted into an ISO/IEC standard. The new standard is for development of other standards. It has already been adopted as a normative standard for European education, in Australia and Canada.

The new Metadata for Learning Resources standard is a multi-part standard and, at the time of writing, is still being extended. While it commenced as work based on the Learning Object Metadata standard in common use world-wide, it evolved into a Resource Description Framework standard in order to maximize its potential for interoperability.

The Dublin Core Metadata Initiative [DCMI]1, as an open community, has collaboratively developed ‘standards’ for twenty years. The deliberately open nature of DCMI work has meant that people with no known connection to DCMI can nevertheless take advantage of the DCMI work and further develop it. This paper asserts that ‘DC Metadata’ has provided a solid base for the MLR and shows, yet again, that DCMI work is thriving in the distributed, global environment.

Keywords: metadata; DCMI; ISO/IEC; Metadata for Learning Resources; ISO/IEC 19788.

1. Introduction

There can be little doubt of the contribution the DCMI community has made to the development of the Web. At a time when it was still thought by many that the Web would fail because nobody would be able to find their way around it, the DCMI rallied the resources of those with information management skills of various kinds and cobbled together a ‘satisfactory’, light-weight cataloguing scheme that showed what could be done with valued resources on the Web. Other ways of finding what would be useful emerged later in the form of search engines using unique algorithms and most recently, referral systems have been shown to be incredibly popular.

While the early selection of vocabulary, representation, documentation, etc. have been honed continually for 20 years, the influence of the early work has extended beyond anyone’s grasp. Today, it is never surprising to find work that has its origins in DC Metadata, wherever it ends up. Significantly, activities and standards such as OAI-PMH, the Semantic Web, the Resource Description Framework, Linked Data, SKOS and schema.org, to name only a few, all have strong connections to DC metadata.

As the publication of yet another derivative DCMI standard, there are features of the Metadata for Learning Resources standard [the MLR] that may be of interest to the DCMI community at large, not only those concerned with learning resources. In particular, the ISO/IEC JTC1 version was written in the context of wide acceptance of the Resource Description Framework [RDF], and with the aim of providing deep interoperability with other educational metadata systems such

1 Where an abbreviation is enclosed as in [DCMI], it denotes a website for which the URI appears in the References section of this paper.
as the IEEE Learning Object Metadata [LOM], benefitting along the way from the work of Nilsson and others (2008).

2. Standards Development

The ISO/IEC Joint Technical Committee 1 follows a standards development process that involves experts and then depends upon votes from ‘participating members’ drawn from the recognized ‘standards’ entities in the participating countries. Only paid-up member countries of the particular ISO/IEC JTC1 committees are entitled to vote and each member country gets one vote but can require as many comments as they make to be resolved. This process, repeated several times in a cycle of editing and voting, leads to the final publication of an ISO/IEC standard. It does not depend on experts in the end but rather national votes. Practice tells us that these votes are not always based on deep knowledge of the issues. This sometimes leads to questions of the wisdom of the decisions, naturally, but the checks are usually in the hands of the experts who do participate in the development of the standards and act as editors for them.

JTC1 standards may be extended as happens when standards prescribe strategies for extension of the standard, as is the case for the creation of additional parts of the MLR standard. Additional data elements can be added to 19788 in new parts, as standards, so long as they comply with Part 1. This is determined on a case-by-case basis by those developing the new part and following the JTC1 process.

DCMI, on the other hand, has maintained a central group of appointed experts to determine the specifications. In fact, DCMI is no longer focused on adding terms to the DC Metadata Terms set but has always encouraged others to develop terms for their purposes and to share them, including by using the DCMI Registry (DCMI 2012). The result has been that significant groups of users now have DC-inspired, and DC-compatible sets of terms that are used in application profiles. In a generic way, it makes sense to talk of ‘DC Metadata’ to mean metadata that is based on DC principles but only those terms published by the DCMI are official DC Metadata Terms [DCMT]. DC Metadata has, over many years now, been formally incorporated into other significant metadata systems, such as the Open Archives Initiative Protocol for Metadata Harvesting [OAI-PMH].

As Wikipedia says:

Dublin Core Metadata can be used for multiple purposes, from simple resource description, to combining metadata vocabularies of different metadata standards, to providing interoperability for metadata vocabularies in the Linked data cloud and Semantic web implementations. (Wikipedia Metadata, 2013)

3. DC Education Metadata

The first DC application profile was built for the Victorian Department of Education in 1999 with the help of Stuart Weibel, the then Director of the DCMI (Nevile, 2008, p 126). It extended the list of DC properties by two, audience and extent (DCMT), giving the education community what it felt it needed for its community in particular, and terms that could be used by anyone else who found them useful.

Typically now, many application profiles take advantage of the extensibility of DC Metadata and include both DCMI recommended terms for global interoperability (DCMT) and local, internally useful terms for local specificity.

Education-related application profiles have proliferated over the years in the broader DC community, and there is no general agreement about what they should include. Sarah Currier’s brief history provides some details:

In 2002 the Institute of Electrical and Electronics Engineers (IEEE) published the IEEE Learning Object Metadata standard (IEEE LOM), superseding the IMS Learning
Resource Meta-data specification, which had been developed and used through several versions since the mid-1990s.

Over the same general period, the Dublin Core Metadata Initiative (DCMI) had established the Dublin Core (DC) as a standard for describing all kinds of web-based resources. The Dublin Core Education Working Group emerged as one of several special interest groups developing specific metadata elements for the use of their communities. (Currier, 2008).

Currier explains how there was a Joint DCMI/IEEE LTSC Task Force that had ambitions of resolving the differences and yet failed to do so because it became “apparent that the LOM and DC had completely incompatible underlying models.” At that time, she continues, the DC community started developing what became the DC Abstract Model [DCAM] and the Task Force worked on expressing LOM using the DCAM. This was not a very satisfactory activity (Currier, 2008-2). The DC-Education community then, Currier said, did some work on an application profile. This work does not appear to have been completed.

The latest news from the DC-Education community is a description of the application profile in development. It has been characterized as a ‘module’ rather than an application profile because it is expected to be used in conjunction with other DC terms and so does not aim to provide a complete set of terms in itself for education. As the DC Education Community website says:

it will only include properties relevant to describing the educational aspects of any resource, and/or the educational context within which it has been or may be used. The intention is that the DC-Education Application Profile will be usable with other application profiles. (DCMI Education Community, 2013)

While it is not clear that there is a ‘sanctioned’ DC application profile for education, such an application profile would probably include some if not all of the first set of fifteen DC terms or those added for the first application profile. Already published as DC Metadata Terms, we now have audience but also mediator, a refinement of audience used to show if a resource is for the student or the teacher, for example; educationalLevel, another refinement of audience; extent, a refinement of format that is used to indicate not the time it takes for a resource to run, as a technical characteristic, for example for a video, but how long it might be used for in an educational setting – eg for a week’s study of a topic at the educational level indicated, and instructionalMethod, intended to be a new term (DCMI Qualifiers, 2005). Both educationalLevel and instructionalMethod are specifically for education whereas the older terms were only included in the extended DC set because they were seen to have other uses beyond education.

4. The LOM

Competing with the rather open set of broader terms for resource description, as described in 2005,

The IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata is an internationally recognised open standard (published by the Institute of Electrical and Electronics Engineers Standards Association) for the description of “learning objects”. The IEEE working group that developed the standard defined learning objects as being “any entity, digital or non-digital, that may be used for learning, education or training”, a definition which has struck many commentators as being rather broad in its scope. … The LOM data model specifies which aspects of a learning object should be described and what vocabularies may be used for these descriptions; it also defines how this data model can be amended by additions or constraints. (Phil Barker, 2005)

It was adopted in many countries for education. The LOM has many terms, specific to education, but many users found that in practice there are too many terms, it is too complicated, and so they only use some of the terms.

The ISO/IEC JTC1 SC36 education community felt the need to resolve the issue and develop what could be at least a single source of properties for their international educational community.
At first this meant adopting the then popular IEEE/IMS LOM metadata schema. As time wore on in the development process, and more people became involved, the need for interoperability was stressed. It was found wanting if the focus was on standardization of the LOM as the new ISO standard but better served if RDF and DC style metadata became the base and the best possible interoperability with the LOM was a priority.

This meant, given the research that had taken place in the meantime, that the new ISO/IEC standard should be fully conformant RDF. There was nothing special about this: DCMI was also working towards harmony with the newer developments, as shown by the work on the DC Abstract Model. DCMI has updated the DCMES and its documentation.

The published MLR is a multi-part ISO/IEC standard that provides first a framework for the creation of properties and their use (Part 1); then adoption of simple DC, the first fifteen elements, with modifications such that the domain for all properties should be ‘learning resource’ rather than just ‘resource’ (Part 2); a basic application profile of the MLR based on the DC properties from Part 2 (Part 3); some technical properties that some felt would be necessary for educational metadata applications (Part 4); and additional data elements based on research of the needs of educators to describe the pedagogical aspects of resources (Part 5); and more ……..


5.1. The Framework (Part 1) and Overview of MLR

As 19788-1 is the Framework for MLR, it does not specify data elements, only how they are to be specified, how Application Profiles are to defined, and so on. This makes Part 1 equivalent, in some ways, to the DC ‘Abstract Model’. The MLR spells out every aspect of data element definition. (Perhaps unfortunately, Part 1 of ISO/IEC 19788 was written before ISO/IEC adopted using online registries for metadata terms rather than having to define them in text only.)

The scope statement (19788-1 s. 1) says:

ISO/IEC 19788 provides data elements for the description of learning resources and resources directly related to learning resources. (ISO/IEC 19788-1, 2010)

This limits the standard to descriptions of learning resources, and Part 2 follows by limiting the domain of the adopted DC Terms but this limitation should be treated as trivial for interoperability purposes (and will probably be removed in the near future).

For a start, a learning resource can be anything. There was contention around the limitation in the adoption of the domain of the terms because that made a different term, technically, and because the use of the terms for resources other than ‘learning resources’ might be prejudiced unnecessarily. The descriptions that will be developed using the MLR, however, are for anything that relates directly or indirectly to a learning resource. This means that within the IT in Learning, Education and Training (ITLET) context, MLR metadata can be used to describe, for example, what is sometimes called meta-metadata (a description of a description of a resource), or a person who contributed to the development of a resource.

The ISO/IEC 19788 scope specifically excludes metadata about a human ‘person’ as a learning resource. Part 9, however, provides data elements (terms) for describing people whose description enriches the description of the learning resource, such as the identity and contact details of the illustrator. In the LOM, such descriptions make use of vCard, an approach also available in MLR Part 9.

5.2. Features of the MLR

It is obvious but interesting that many resources become learning resources well after their first publication. Most resources have titles, subjects, coverage, and more, whatever they are expected to be used for. So the MLR team chose to work with already widely adopted general metadata, thus enhancing the opportunities for interoperability (MLR-1, 2012, Clause 5.5).
The ISO/IEC 19788-1, the Framework, is for two classes of users: developers of other parts of the standard itself, or other standards, and independent communities that develop terms for themselves or others to use.

As Gilles Gautier has said (2013), the MLR has a very different approach to names of data elements from other systems, such as the LOM or DCMT. In the MLR, the data element identifiers are linguistically neutral (non-linguistic) and it is left to each National Body to decide what labels to use in their language and writing system. Gauthier (2013), a native French speaker, points out that despite a term being associated with a string such as "creator", to the computer this should be only a string of letters, and for humans there should be a facility for names in multiple languages. DCMI has wrestled with the problems of internationalization for many years, as have others.

The MLR approach supports internationalisation by allowing any country to provide the name/labels in their own language. Every one is able to use their own language for the properties and classes and the global interoperability (if needed) is provided by making available, for example in the form of a SKOS vocabulary, the equivalencies on the Web. The aim is local diversity while maintaining global interoperability.

The MLR terms (data elements) will soon be released with online representations of the names in multiple languages in a system that will provide for as many languages as are submitted (as a result of simultaneous revision work by ISO/IEC JTC1 SC36 WG 1 updating ISO/IEC 2382-36).

5.3. Terms

'Terms' as referred to by DCMI are ‘data elements’ in the MLR:

A data element is a 3 or 4 parts entity, either

<dataElementSpecificationID, subject, contentValue>

or

<dataElementSpecificationID, subject, contentValue, languageCode> (Clause 7.1, 19788-1)

The data element, or terms specification in Part 1 has adopted DCMI practice. That is, the table that was used by DCMI for definition of terms has become part of the MLR standards specification process.

When DC Metadata was first developed, the community did not see the need for a strict definition of domains and ranges. Later, given the intended scope of DC Metadata, it was assumed that the domain would always be resource, unconstrained, so this has now been included in the definition of the DCMT. The range has proved more problematic and is now formally stated for each term. For today’s technology, it is important but in the early days of the Web it was accepted that a URI could be a literal. The MLR does not endorse this practice.

5.4. MLR-2 adoption of DC Terms and Values

MLR has modified the DC definition of the data element refines and expands this with a rule to ensure that controlled vocabularies, as the DC community calls them, are not broken by the addition of new terms. Here, while it is hoped a community using DC Metadata will not add a term to a vocabulary list that causes confusion, the MLR makes it more difficult to do this. The MLR constraint is that it is conformant to replace one vocabulary with another in an application profile only if:

the data element is repeatable with the condition that: whenever a value from the extended or complementary part of the vocabulary is used, a value from the original vocabulary is also used (19788-1, Clause 12.4, Rule 0050)

This rule is already targeted for replacement by a simpler solution in the next edition of the standard. (ISO/IEC JTC1 WG4 private discussions, 2013)
DCMI does not impose ‘rules’ but has precedents from community practice such as the ‘dumb-down’ rule. This rule supports backward compatibility and interoperability and is used to ensure a new term in a values vocabulary, for instance, does not break interoperability. A new term must be a refinement of an established term, or completely distinct from all others. For the DCMI community, adding the term moving image to the list of values including image for type of resources was contentious because it threatened to break this rule.

5.5. MLR Application Profiles

As an MLR application profile can include data elements from other sources, users may opt for the adoption of original DCMTs establishing a mixture of them and MLR data elements, or possibly just DCMTs. From the DC Metadata perspective, this is also acceptable. Just how interoperable or conformant such an application profile will be to the DC model is perhaps best determined by reference to the DC Singapore Framework.

5.6. MLR Data Element Identifiers

Identifiers have evolved considerably since the early days of DC Metadata. ISO/IEC obliged the editors of the MLR to include identifiers in the standard, with verbal definitions. These are clumsy, by comparison with registry entries that would be allowed now. Nevertheless, the editors worked to provide unique identifiers that could be published online and it is possible to resolve them by reference to a namespace following a complicated set of rules for the definition of the terms (see ISO/IEC 19788-1, footnote to 6.2.1.1, Rule 004).

5.7. MLR Records

A MLR Record is a specified set of data elements for describing a specific learning resource and resources directly related to that learning resource. (19788-1, Clause 10.1)

This replicates the DC 1:1 rule and Abstract Model in that a description is to be of a resource directly, with the domain specified, but there can be a record with several of these included. An
MLR Record is the equivalent of a DC Description Set. In practice, descriptions might be of resources that are related and can be ‘chained’ (as the author thinks of it). An example is where the personal details of an illustrator of a learning resource are being described. Here the links run from the resource to the person and then to the description of the person. Each link in the chain has a subject (legitimate according to the domain of the term) and a value (within the term’s range) that becomes the subject (again legitimately within the term’s domain) for the next value.

5.8. MLR Part 5 – Educational Data Elements

The ISO/IEC JTC1 community worked with others to develop a concept map of their needs. The resulting diagram (Figure 2), illustrates the data element specifications for 19788-5, the part where educational, or pedagogical data elements have been added to 19788.

As can be seen in Figure 2, the domain for some of the data elements introduced in Part 5 varies according to the data element. There is a domain of contribution, and this is described by a number of properties, as is the potential audience for the learning resource. This is not the approach that DC Metadata usually takes, where a resource is described by a property or a refinement of that property. In the case of DC Metadata, a lot of the elegance of the metadata has come from the technique of using refinements.

![Figure 2: Diagram of data element specifications for ISO/IEC 19788-5 (Appendix C of Part 5)](image)

A data element such as annotationDate could describe an annotation to the original resource and be a refinement of date, following the DC Abstract Model. The advantage of this approach,
of course, is that the DC ‘dumb-down’ rule would apply. The advantage of working the MLR way is that it is RDF compliant and there are tools, such as SPARQL, that would search and find the metadata.

Personally, the author favours the DC approach for elegance and simplicity. There are fewer things to think about! The structure is more obvious, in some ways. An annotation can be a significant resource, in itself. In fact, it can be described in the same way as the other resource, and simply related to it, using the relation data element from Part 2.

It is suspected that the way Part 5 has been conceived has come from LOM thinking, where hierarchical sets of metadata were common. The MLR aims to flatten everything. So an interesting question might be, “What about a crosswalk from the MLR to DC Metadata, or the reverse?” Not all the properties considered by the MLR have equivalent DC terms, of course. Another might be, “What if an ‘annotation’ is a learning resource? Does it have two sets of descriptive metadata? or a combination of both? It is easy to see how one goes from a learning resource to an annotation but how messy will the map become when the annotation, as a learning resource, has an annotation?

5.9. MLR Interoperability

MLR Part 3 offers an informative crosswalk from the LOM to the MLR (19788-3, Annex C1). It was a long and painful process shifting the emphasis from the LOM to the MLR, changing from databases and records to triple stores and RDF. Now that the MLR is mature, and the change has been supported, there is still cleaning up to do to some of the definitions to ensure the switch is complete. There are also areas for which more data elements are expected. There is a revision cycle for such standards and it is also possible there will be a second version once there is more experience with the use of the MLR.

A DC (education) to MLR crosswalk is shown in ISO/IEC 19788-3 Informative Annex C2:

<table>
<thead>
<tr>
<th>Dublin Core</th>
<th>ISO_IEC Identifier</th>
<th>MLR</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>audience</td>
<td>19788-5:2012::DES0400</td>
<td></td>
<td>audience language</td>
</tr>
<tr>
<td>educationLevel</td>
<td>19788-5:2012::DES0500</td>
<td></td>
<td>audience level</td>
</tr>
<tr>
<td>instructionalMethod</td>
<td>19788-5:2012::DES0600</td>
<td></td>
<td>learning method</td>
</tr>
<tr>
<td>mediator</td>
<td>19788-5:2012::DES0700</td>
<td></td>
<td>audience role</td>
</tr>
<tr>
<td>subject</td>
<td>19788-5:2012::DES1200</td>
<td></td>
<td>curriculum topic</td>
</tr>
</tbody>
</table>

There are also crosswalks provided in the annexes of 19788-5. For the future, as a matter of policy, SC36 has decided to provide information about how to transform LOM metadata into MLR-conformant metadata rather than trying to maintain the two in parallel.

6. The Diagrams from the MLR

The MLR contains a number of diagrams designed to help with its understanding and use. Many of these may also be useful to DC users.

In developing the MLR in its current form, it has become clear that, for many, changing from databases and records to triple stores and RDF involves a difficult ‘leap of faith’. This problem is familiar in the DC world where some say the DC Abstract Model helps and others find it totally
confusing: some people find diagrams difficult, and prefer words, while others work better with diagrams.

In particular, as users are not expected to have to understand the definition of data elements in the MLR, or terms in the DC Abstract Model, the problem is usually only for developers of new data elements. The use of diagrams has clarified, if not made understandable, how the data elements should be developed. Where there is confusion about the domain, for example, there is a need to work more until it is clear that the domain of a data element is as it should be. This work is often best done by working dynamically with RDF diagrams. Particularly the process of working out what is a unique ‘class’ and what is a refinement or in conflict with another ‘class’ is often made clear by diagrams.

As an example, the concept map for 19788-5 might have led to a different set of data elements. If the resource class was not different for, say, an annotation, there might be two resources for description that would both be thought of as ‘learning resources’. Then the description of the annotation might use many of the same terms as are used for the original learning resource with the difference that it would be identified as in relation to the original as a resource of type annotation. These latter terms could then be standard DC Metadata Terms, or MLR adaptations of DC terms (19788-2). Characteristics such as the date of the annotation could appear, sensibly, in two places; once as the date on which the resource was annotated (with subject of learning resource) and secondly as the date of creation of the resource that is the annotation (with subject annotation). The first, the date of annotation as a property of the learning resource, would be a refinement of the term date if this were being done according to the DC Abstract Model and could be, according to the MLR Framework.

The way that 19788-5 defines the metadata now, if a learning resource is an annotation of another learning resource, it will not be described using the same data elements as the original which might lead to confusion.

7. Conclusion

It is expected the ISO/IEC 19788 standard for Metadata for Learning Resources will be used widely within education around the world. As well as formal adoptions in some countries, significant participation in its development by experts from China supports this expectation. As education becomes a global enterprise, this is important for teaching and learning users.

Perhaps it is the case that what has previously been known as ‘qualified Dublin Core’ has been adopted by ISO/IEC? If the Framework of 19788-1 establishes the equivalent of the DC Abstract Model, and then adopts the DC terms, not just with modifications as in Part 2 but exactly and completely in another part, or in an Application Profile, arguably the result will be DC Metadata.

Rather than be concerned about what is or isn’t DC Metadata, this paper shows that outside of the DCMI, people who have worked in the DC community or with DC Metadata are busily developing new versions of it and standards communities are building fruitfully on the earlier DCMI work. This should be considered as yet another accomplishment of the DC community.

Acknowledgements

The work of Professors Gilles Gauthier and Yolaine Bourda, in particular, are recognized along with the contributions of many other members of the ISO/IEC JTC1 SC 36 WG 4, for sharing the learning and development of the Metadata for Learning Resources standard ISO/IEC 19788. Professor Gauthier and Charles Nevile offered very helpful comments on this paper.

Abbreviations and Links

DCMES: Dublin Core Metadata Element Set (ISO/IEC 15836)
DCAM DCMI Abstract Model ???
References


