
Metadata in the changing learning environment: developing skills to achieve the blue skies

Juanita Foster-Jones and Hazel Beazleigh
The Open University
email: J.Foster-Jones@open.ac.uk

This short paper will examine the importance of metadata and its role in the changing learning environment, beginning with an introduction about what metadata is, and the benefits to be gained from applying it to all academic resources. Two Open University projects, Portfolio and the Reusable Educational Software Library, will be described and used to illustrate how the IMS Learning Resource Metadata scheme is being applied, and the issues that have been encountered by the Open University and how it is attempting to resolve them. The need for change in organizational culture so that metadata becomes part of the creation process, rather than an afterthought, will then be discussed. The paper concludes with a glimpse into the blue skies of the future – where all resources will have metadata as standard practice, and institutions can share and utilize their resources effectively.

Defining metadata

It is perhaps a reflection of the increasing importance of metadata that the term is now commonplace amongst publishing, computing and library and information communities and, as such, subject to many attempts to define it properly. Definitions range from simplistic attempts such as, 'stuff about stuff' (Dovey, 2000), to more comprehensive efforts to capture its meaning and purpose fully. Of these, the definition we feel is most enlightening states 'metadata describes an information resource. It is the internet-age term for information that librarians traditionally have to put into catalogues, and it most commonly refers to descriptive information about Web resources' (Hillman, 2000).

The underlying concept of metadata is not an unfamiliar one. Descriptive information has long been applied to resources to enrich their existence and more recent metadata developments simply reflect our need to manage the proliferation of electronic resources with which we now have to cope.

Metadata in the changing learning environment

As technology evolves, so the landscape of learning is clearly changing. No longer confined to the school, further education, and university education progression, learners can choose when and how they acquire knowledge, and lifelong and modular learning have become commonly sought after educational experiences. The Open University (OU) has always offered its students a flexible approach to learning, so it is perhaps unsurprising that we are now also embracing ways to extend our expertise into new realms of information delivery.

Our challenge is to meet users' expectations of flexibility and ease of use, prompted by their experiences of many online environments; by facilitating access to information, so that their efforts may be concentrated on achieving individual objectives, rather than merely searching for the objects to achieve them. We are convinced at the Open University that effective resource description could indeed provide the key to meeting this challenge by:

- enabling users to discover and identify existing materials;
- providing them with sufficient data to evaluate and distinguish between different resources;
- providing the option to personalize information presented, via a learner's information profile.

Open University case studies

We will use two OU projects to illustrate the issues raised while implementing metadata, Portfolio and the Reusable Educational Software Library (RESL),¹ which, although they differ significantly in their implementation, share the following features:

- the use of IMS Learning Resource Metadata Specification v1.0²;
- the objective of facilitating reuse of existing materials;
- a change in responsibility for metadata creation from the hands of experts (information professionals) to those of the practitioners (content creators).

Portfolio digital archive

The ultimate function of the Portfolio digital archive is to provide a repository for storing, locating and reusing print-masters of all OU course materials. To achieve this, we have adopted a phased metadata implementation strategy, selecting in the first instance a skeleton set of IMS metadata elements (see Figure 1) in order to test their impact on locating and evaluating records. This was driven by the size of the database (10,000 items initially with 500 items added per month) and the resource implications. We felt that it was more likely to be successful if a small number of elements were chosen. Our test results will then be used to guide subsequent phases of development, which will focus on how users may select items for reuse on the basis of pedagogical rationale, using educational elements.

Currently, metadata is being added retrospectively to each item within the archive, which is a huge and possibly insurmountable task, as approximately 500 new items are being added to the archive every month. Work is therefore in progress to develop tools that will facilitate

Portfolio Metadata	
Metametadata Metadata Agent Role <input type="text"/> Metadata Agent Name <input type="text"/> Metadata Agent Contact <input type="text"/> Date of Metadata Creation <input type="text"/>	
General Elements Item Full Title <input type="text"/> Catalogue Name <input type="text"/> Catalogue Entry <input type="text"/> Item Description <input type="text"/> Keywords <input type="text"/>	
Lifecycle Elements Document Version <input type="text"/> Contributor Role <input type="text"/> Contributor Name <input type="text"/> Contributor Contact <input type="text"/> Publication Date <input type="text"/>	
Educational Elements Learning Resource Type <input type="text"/> Intended User <input type="text"/> Educational Level <input type="text"/> Description <input type="text"/>	
Classification Elements Classification Scheme <input type="text"/> Keywords for Classification <input type="text"/> Broad Subject Category <input type="text"/>	
Relation Elements Relation Type <input type="text"/> Related Resource <input type="text"/> Relation Type <input type="text"/> Related Resource <input type="text"/> Relation Type <input type="text"/> Related Resource <input type="text"/> <input type="button" value="Add New Resource"/> <input type="button" value="Save"/>	

Figure 1: Portfolio Metadata Elements – Phase 1: IMS descriptive elements have been used for Portfolio, in conjunction with other University projects to secure its future interoperability with other repositories

concurrent, author-generated metadata – a goal shared by other projects within the OU – within the course production workflow. This will ensure that authors will not be burdened with additional work, but that metadata they already generate is systematically collected.

RESL

The Reusable Educational Software Library (RESL) is a strand of the SoURCE (Software Use, Reuse and Customization in Education) TLTP project.³ RESL originally aimed to test the feasibility of a national software library through the creation of a prototype and encourage academics to add examples of software to the prototype. This concept was later expanded to encompass case studies created by SoURCE and other TLTP projects in order to enhance the feasibility study, whilst facilitating the dissemination of information. Potential users of RESL were identified as software developers, academics and administrators, who would each have different needs. These were investigated in RESL-OU-1 (available from RESL), and resulted in the use of four different controlled vocabularies, with terminology aimed at satisfying the different information needs.

It was assumed that academics would have little experience of metadata creation, so data entry templates were created. These were converted into data entry forms, presenting the metadata author with only the fields required for a specific type of resource to ensure consistency and quality of metadata supplied. A final proofing stage was also built into the workflow as part of the quality control process. The use of controlled vocabularies, represented as 'pick lists' in the data entry templates, maintained authority control, whilst free text keywords allowed a degree of flexibility for the metadata submitter. The controlled subject vocabularies are also used as browse lists on the search interface, linking metadata to the information retrieval process. This recognizes the fact that searchers may not have a clear idea of what it is they are looking for, or be able to articulate it (Belkin, 1982). By providing browse lists, searchers have a starting-point to guide them to the resources, which will, we hope, lead to a match for their needs.

RESL implements selected elements from all of the IMS categories. Additional project-specific metadata elements have also been added to fulfil certain project needs, but only after much deliberation since they could impinge upon interoperability. Interoperability was seen as highly desirable; if RESL becomes a national database we would wish to share records with other sources. Where IMS was extended, it was only after we had checked that no other field was suitable for our purpose, and only if we felt that the field would give benefit to the users, or increase the functionality of the database. For example, we added a 'date last visited' field after the URL, so that users had an idea of the currency of the information provided. We also added a project field to the metametadata screen, allowing projects to submit their project URL and logo, thus identifying resources they had submitted. It was felt that this would encourage other TLTP projects to add resources. It should also be noted that RESL does not use XML binding for its metadata, due to time and resource constraints.⁴ If the project were to be scaled, then XML would need to be considered to ensure interoperability with other IMS compliant systems.

Issues

For both projects, the primary concern was to determine whether to use an existing metadata schema, or to create an OU-specific schema. In both instances we decided to

adopt an existing schema. This ensured that time was not wasted in duplicating effort, and would facilitate future interoperability. IMS was chosen for its semantic density, in particular for the educational elements. The OU also has a responsibility as joint partners in CETIS to contribute to IMS development and thus, by implementing our metadata, we can then feed our experiences into future revisions. However, adoption of this schema has identified a number of issues.

The IMS metadata model consists of nine categories, with a total of sixty element tags for describing a learning resource, all of which are optional. Having decided upon using IMS, we then needed to select which elements to implement. Different projects have different needs, and so will use a variety of element combinations. In addition, we also needed to decide on the depth of description to give. For Portfolio, a staged approach was taken, starting with a minimum number of elements with the option to increase as required, whereas for RESL a comprehensive approach was taken, implementing a large number of elements to test the IMS metadata. Portfolio used the staged approach for two reasons. Firstly, due to the sheer size of the database, choosing a limited number of elements to complete would ensure more items were given metadata. Secondly, Portfolio was investigating the levels of granularity to which metadata should be added, that is, should metadata be added at the item, chapter or paragraph level. A small set of metadata would facilitate this.

For RESL we wanted to take advantage of the semantic density of the IMS schema, to facilitate in resource identification and evaluation. For example, in the technical category, IMS provides a number of elements which describe the hardware and software requirements of the resource. These were felt to be of use to the end user, as they would provide an indication of what equipment is required to access the learning resources described on the database. Institutions will need to decide what level of metadata entry they can resource.

Selecting appropriate elements has proved to be no simple task. In some cases this has been hampered by the lack of examples provided with IMS specifications. Our experience has shown that descriptions within the guidelines can be complex and difficult to interpret, due to their terminology and technical bias. Although we were approaching this as qualified librarians, and thus as experts in metadata, we have still experienced difficulties. To this end we have decided to look at all metadata elements within the IMS specification, and give examples of how they may be used within the OU context. This will make adoption by other projects easier, and encourage a standardized implementation.

Where IMS specifications refer to other documents for best practice it can be difficult to locate the relevant part, as you are referred to a main site. Some of the documents referred to are also aimed at users with technical backgrounds, and can be complex to read and understand, often hindering progress.

It should also be noted that IMS did not meet all requirements of either project, although we recognize that it is unlikely that any standard would accomplish this. Both Portfolio and RESL have had to make decisions about whether to add metadata elements to the schema. This was addressed with caution, on the understanding that any deviance would impinge on interoperability. Where additional fields were added, it was only after all IMS elements had been considered for their suitability. Guidelines within the IMS Meta-Data Best

Practice and Implementation Guide (Anderson and Wason, 2000) on extending IMS would have been welcome.

Finally there is the fact that IMS is a specification, and hence susceptible to revision. A plan of action needs to be determined to review each new version, and decide to what extent any changes should be implemented. It may be that metadata is upgraded for new items, but those using the old schema are left unrevised due to lack of resources.

Lessons learnt

The OU is still in the early stages of metadata implementation. As we become more experienced in our use of metadata, we shall be able to address some of the issues raised by these two projects. As further projects recognize the importance of metadata, it is likely that new issues will be raised. There is still a considerable amount of work to be done to ensure that metadata is created and exploited to its full potential, and that the issues identified above are overcome.

Experience from Portfolio has shown that retrospective metadata application to existing, populated databanks is a Herculean task. Methods of creating metadata automatically at the point of item creation would help address this but would require working practices to change, to make metadata generation an integral aspect of authorship. Tools such as DC-dot⁵ and the Sun Microsystems Developers toolkit⁶ can assist in this automatic generation of metadata, and work will need to be carried out to investigate their effectiveness, and then compare the quality of metadata generated by tools such as these with manually generated metadata.

For any such change to be successful it requires authors to become stakeholders, and to share ownership of metadata. To do this we must first prove its value, by making resource location more effective and by removing fears of extra work by automating the process as much as possible. The OU is currently examining the use of XML style-sheets to see if they can provide a workable solution.

Metadata needs to be an integral part of an institution's information strategy. Recognition and support for metadata is required from the top of an organization's structure. Only with official recognition will it be possible to put measures in place for a standardized approach to metadata creation and quality control. Within the OU a Metadata Working Group has been created, to provide recommendations for a metadata strategy. This will be presented to OU boards for approval, and by so doing will set metadata firmly on the OU agenda.

There is also a need for implementers to share implementation and best practice examples, and to build communities of practice. Given that many metadata schemas are not yet standards, there is an opportunity for early adopters to influence their further development.

Blue skies

It is too easy to become distracted by the obstacles and issues to be overcome, and to question the relevance of metadata. But it is crucial that we invest our efforts and resources in this developmental phase so that we can contribute to the evolution of metadata standards, using our experiences to create one that best suits our needs.

We should not lose sight of the ultimate objectives of this exercise, which are to locate, share and reuse the learning resources that are the assets of our organization. Like information, a resource is only valuable if it can be found and utilized in time. 'The value of information often depends on how easily it can be found, retrieved, accessed, filtered and managed' (Martinez, 2001).

Perhaps the way forward to achieving interoperability will be through the use of application profiles. Application profiles 'consist of data elements drawn from one or more namespace schemas combined together by implementers and optimised for a particular local application' (Heery and Patel, 2000). In the blue skies of the future there will be a registry in which application profiles are stored. These will be accessible to implementers, so that they can benefit from work already completed, and to creators of metadata tools, to assist with the development of metadata creation and information retrieval processes that take into account this 'pick and mix' approach to metadata.

Utilization of the Resource Description Framework (RDF)⁷ will contribute to this tailored metadata approach:

The Objective of RDF is to support the interoperability of metadata. RDF allows descriptions of Web resources – any object with a Uniform Resource Identifier (URI) as its address – to be made available in machine understandable form. This enables the semantics of objects to be expressible and exploitable. Once highly deployed, this will enable services to develop processing rules for automated decision-making about Web resources. (Iannella, 1998)

RDF allows a resource to be described using elements from multiple schemas, in effect providing the structure for application profiles.

The aim of all metadata is to make resource location possible. We value libraries for structuring and managing books for research, therefore we should no less value metadata creators for structuring and managing our learning resource assets. At the OU we are making a commitment to metadata standards, to ensure the future accessibility of all our assets.

Notes

- ¹ RESL: <http://www.resl.ac.uk/> (last accessed 17 December 2001).
- ² IMS: <http://www.imsproject.org/index.html> (last accessed 2 January 2002).
- ³ SoURCE: <http://www.source.ac.uk/> (last accessed 4 October 2001).
- ⁴ A full list of RESL metadata elements, and associated rationale can be found in RESL-OU-4 and RESL-OU-5 on the SoURCE Web site <http://www.source.ac.uk/>.
- ⁵ DC-dot: <http://www.ukoln.ac.uk/metadata/dcdot/> (last accessed 17 December 2001).
- ⁶ Sun Microsystems Developers Toolkit: <http://www.imsproject.org/tools/sun.html> (last accessed 17 December 2001).
- ⁷ RDF: <http://www.w3.org/RDF/> (last accessed 2 January 2002).

References

Anderson, T. and Wason, T. (2000), 'IMS Meta-Data Best Practice and Implementation Guide', version 1.1. Available from: <http://www.imsproject.org> (last accessed 25 October 2001).

-
- Belkin, N. J., Oddy, R. N. and Brooks, H. M. (1982), 'ASK for Information Retrieval: Part 1, background and theory', *Journal of Documentation*, 38 (2), 61–71.
- Dovey, M. J. (2000), '“Stuff” about “stuff” – the differing meanings of “metadata”', *Vine*, 116, 6–13.
- Heery, R. and Patel, M. (2000), 'Application profiles: mixing and matching metadata schemas', *Ariadne*, 25, <http://www.ariadne.ac.uk/issue25/app-profiles/intro.html> (last accessed 26 July 2001).
- Hillman, D. (2000), 'Using Dublin Core'. Available from: <http://dublincore.org/documents/2000/07/16/usageguide/> (last accessed 18 April 2001).
- Ianella, R. (1998), 'An idiot's guide to the resource description framework'. Available from: <http://archive.disc.edu.au/RDUIreports/RDF-Idiot/> (last accessed 19 December 2001).
- Kim, Y., Clarke, S. and Ramsden, A. (1999), 'Baseline report setting the context for Reusable Educational Software Library (RESL) of SoURCE Project (RESL-OU-1)'. Available from: <http://www.eres.ac.uk/source/docs/resl-ou-1.pdf> (last accessed 17 December 2001).
- Martinez, José M. (2001), 'Overview of the MPEG-7 Standard', ISO IIEC JTC1/SC29/WG11 N4031. Available from: <http://www.csel.it/mpeg/standards/mpeg-7/mpeg-7.htm> (last accessed 16 May 2001).