Registry for Educational Metadata Schemas

A proposal to the JISC and BECTa from UKOLN and the ILRT

Document Notes

Author	Rachel Heery, Dave Beckett
Date	19 February 2002
Version	Final version
Document Name	reg-jisc-bid20020219.doc
Notes	This document has been prepared for consideration by the Joint Information Systems Committee, in particular the JISC DNER Office, and BECTa.

1.   Executive Summary

This document provides a proposal for the development of software tools for the prototype schema registry established by the Metadata for Education Group (MEG). The aim of the MEG registry is to provide implementers of educational systems with a means to share information about their metadata schemas and to re-use existing schemas. The benefit will be a saving of time and effort currently spent in researching existing schemas and in re-inventing schemas.

The proposal is to develop and enhance the existing software underlying the MEG registry (based on demonstrator software developed during the EC funded DESIRE project) in order to move towards a sustainable and scaleable solution. The central registry will be hosted at UKOLN. Enhancements in software will focus on

- creation of machine-readable schemas by schema creators (in this instance members of the MEG group) so that they can be input to the registry without the intervention of a human administrator rather which is a severe limitation within the existing system

- use of the Resource Discovery Framework (RDF) schema language for input and output of schemas in order to facilitate interoperable exchange of schema data between systems

This proposal includes a workshop to support MEG members in the process of inputting their schemas in a structured format using the new schema creation tool. This will provide the end result of a Registry populated with schemas.

The proposal is made by UKOLN with software development to be undertaken by the ILRT. The proposal builds on UKOLN's experience in designing schema registries (within the DESIRE and SCHEMAS projects), and UKOLN's knowledge of the requirements of the MEG group. The proposal builds on the ILRT's expertise in software development and their experience of development work using RDF/XML technologies. Both organisations have experience of dissemination and liaison activities, as well as close links with the stakeholder communities, and will be well placed to run a joint workshop for members of the MEG group.

Development of a schema registry for the MEG group is seen as a means to inform further work on schema registries in the context of the developing Information Environment. UKOLN and the ILRT are closely involved in DNER developments and will be in a position to consider how schema registries might play a role within the wider context of the Information Environment.

2.   Benefits of metadata schema registries

Web based services use structured information to describe characteristics and content of web resources in order to manage those resources, whether this is in the context of descriptive metadata, terms and conditions, preservation, etc. Such structured information (metadata) exists in many different element sets (schemas) such as EAD, MARC, IMS, DC, NewsML etc. This variety of schemas will no doubt continue in order to support specialised systems designed for particular uses. This is the case even within domains. So, for example, within the learning technology domain there is use of DC-Education and IMS. However, there is a further level of difference. Particular communities of use may propose 'extensions' or 'variations' to a particular standard to cater for specific requirements.

Such differentiation brings a burden in terms of interoperability between systems. Certainly unnecessary variations in schemas should be avoided wherever possible, but it is impossible, and undesirable, to try to impose complete uniformity. What this means is that there is a growing need for users to find out about schemas in use, and about particular extensions or variations. Currently, these 'users' are human – people would like to look something up. Increasingly, such 'users' will be automated – 'agents' which hide some of the difficulty of translating between formats from human users.

Research within the SCHEMAS project, and experience within the DCMI has led to development of schema registries to facilitate

- Increased interoperability between schemas

- Less duplication of effort

- Promotion of existing solutions

3.   Background to MEG Registry

Currently there are over sixty members of the MEG group with approximately twenty known to be active in creating schemas to describe educational resources. UKOLN acts as co-ordinator for the group. The existence of such a focused group offers great potential for sharing and collaboration regarding design and re-use of schemas.

Consensus has been reached within MEG that a registry is required to

- Provide easy access to standard educational metadata schemas through a Web interface which allows the user to navigate relationships between the components of those schemas

- Provide information about extensions to standard formats used by MEG members

- Give authoritative definitions for local elements, and permitted schemes for elements

Users of the MEG registry would typically come under the following categories:

- Publishers of standards

- Implementers seeking appropriate   schemas

- Developers comparing schemas

- Metadata creators seeking assistance in using particular schemas correctly

So, the aim is that, in the UK learning space, the MEG registry will provide a mechanism for various learning initiatives to manage interoperability between their metadata implementations by publishing their schemas in a registry which is available to others. This in turn will provide a concrete way of

encouraging sensible uniformity alongside necessary divergence. It will help avoid unnecessary recreation of wheels, and support sharing of approaches. Experience suggests it is a more viable approach than trying to impose a 'one size fits all' message.

The work on registries within the DESIRE project (which included ILRT and UKOLN as partners) and the SCHEMAS project (which included UKOLN as a partner) was recognised as being relevant to requirements for sharing information about metadata schemas amongst the several educational initiatives within the UK. In addition experience in implementing registries has been gained within the context of the DCMI Registry Working Group which Rachel Heery from UKOLN co-chairs.

This experience enabled UKOLN to establish an initial registry prototype within the MEG context, providing access to educational related metadata schemas. UKOLN has created entries for some metadata schemas in this registry but the system is based on demonstrator software developed within the DESIRE project and enhancements are urgently required to deliver a more sustainable, maintainable and scaleable solution.

We hope that as a first step towards an operational service we can enhance the software tools on which the registry is based. This would require

- ensuring the data model is extensible

- improving ease of data input to enable distributed input of schemas by MEG members

4. Deliverables

4.1 Review of existing software solutions

This will be undertaken by ILRT who will review current software options for registry implementation with a view to identifying suitable existing software. A summary report of findings will be made available linked from the MEG web site.

4.2 Functional specification

This will be undertaken by UKOLN in collaboration with MEG users. It will focus in particular on a revised data model for the MEG registry and on new areas of functionality required for the schema registration tool.

4.3 New implementation of registry based on sustainable software

This implementation would have similar functionality to the existing MEG Registry. This will be undertaken by ILRT and will involve deploying a new lightweight registry based on the current MEG registry with some of the existing functionality surplus to MEG requirements cut out. It will allow the loading of schemas, and the re-loading of updated schemas, browsing and querying the registry for terms, schemas (in namespaces) and application profiles.   It will take advantage of relevant technologies such as XML and RDF schemas. The aim is to provide a platform for a sustainable, easily maintained and scaleable and existing web technologies will be used.

4.4 Schema registration tool

Development of the tool will be undertaken by ILRT according to the functional specification produced as in (4.2) by UKOLN. Currently all schemas are input to the prototype MEG Registry by UKOLN staff. This is not scaleable and is a major barrier to growth in the number of schemas being registered. A tool will be developed that will enable distributed creation of schemas which can then be input into the registry automatically. The intention is that the tool will enable the schemas to be output from the creation tool in a standard format (probably as RDF schemas). The tool will allow for the creation of schemas based on standards but with local extensions (application profiles).

4.5   Workshop for MEG users to facilitate them in populating registry

A joint MEG workshop will be run by UKOLN and ILRT (scheduled for July if possible) where the new registration tool will be introduced to MEG members. MEG members will be supported in using the tool to create their own schemas in a standard format to enable them to be infused into the MEG registry.

4.6   Recommendations for future

Experience from the MEG project will be reviewed by both ILRT and UKOLN and high level recommendations will be made in summary form regarding the future role of registries in the DNER context.

5.   Looking ahead

This proposal will result in enhanced tools for running a registry. Following completion of this work it would be worthwhile to undertake user trials of the registry, validating the design and investigating the usability of the registry. Such trials would inform establishing an operational service. Additional funding would need to be sought for this work.

Metadata registries may in the future provide additional 'added value' services in the Common Information Environment. As well as providing information about schemas, registries might also provide mappings between schemas, links to good practice guidelines, contacts and reference implementation information, even background on standards and associated protocols.

The ambition motivating the proposal is to create a sustainable 'publishing environment' for metadata schemas which would form the basis for extended functionality in future. In addition it is hoped that the initial work on a registry for educational metadata schemas would form the basis for a future operational MEG registry service and, if appropriate, a generic registry service which might become a component of the DNER Information environment.

Future enhancements in functionality (not covered by this proposal) might enable the registry to

- Provide mappings between the terms of different element sets

- Annotate elements (e.g. by status such as proposed new element or with user guidance etc)

6.   Project partners

6.1   UKOLN

UKOLN is a national focus of expertise in digital information management. It provides policy, research and awareness services to the UK HE and FE, information and cultural heritage communities. UKOLN hosts three of the JISC-funded Focus posts and has a strong culture of dissemination, organising and participation of events and liaison activities. This work is complemented by UKOLN's research and development activities and the provision of a range of national and international services. UKOLN's key strengths of relevance to this bid include expertise in metadata, schema registries and resource discovery

UKOLN's liaises with and supports services and projects within the education sector. The work of the Interoperabilty Focus within MEG is particularly relevant to this proposal. UKOLN has undertaken joint project work with the ILRT within JISC and EU projects and has many informal links with ILRT which will help ensure success in this work.

6.2   ILRT

The ILRT is a centre of excellence in the development and use of learning technologies and is host to a number of JISC funded services and projects. The ILRT has a wealth of experience in managing projects, dissemination activities, developing and facilitating training workshops, project evaluation and the development of teaching and learning materials.

With respect to this bid, key ILRT strengths include their research and development work in XML and RDF technologies, and close links with standards making activity in W3C RDF initiative. ILRT are experienced in software development of web based tools and in running software development projects.

7.   Management

7.1   Project management and financial arrangements

UKOLN will act as co-ordinators and project managers for the work, and will act as main contractor to the JISC. UKOLN will provide liaison with MEG and with the DNER team. Development work will be undertaken by the ILRT in the role of subcontractor (given the scale of the project this will avoid unnecessary bureaucracy). Financial arrangements will be agreed between ILRT and UKOLN.

A project team will be formed consisting of two people from UKOLN (Rachel Heery and Pete Johnston) and two from ILRT (Dave Beckett and one other).

UKOLN will produce a summary project plan which will outline responsibilities, dates of deliverables, format and frequency of meeting and financial arrangements.

7.2   Getting Started

We are looking to start work on this project in March and intend to complete work over an elapsed time of six months.

7.3   Software and IPR

All software developed during the project will be developed from existing open source software (also called free software) and will be made freely available under similar or equal-freedom licences as defined at http://www.opensource.org/licenses/

Credit will be given to both institutions in all deliverables including software, papers and reports.

8.   Useful Links

DESIRE Registry: http://desire.ukoln.ac.uk/registry/index.php3

MEG Registry: http://www.ukoln.ac.uk/metadata/education/registry/contents.html