Interoperability issues between learning object repositories and metadata harvesters

Ricard de la Vega¹, Jordi Conesa², Julià Minguillón²

¹Centre de Supercomputació de Catalunya (CESCA)
²Universitat Oberta de Catalunya (UOC)

Open Repositories, July 6th-9th 2010, Madrid, Spain
Table of contents

Introduction
UOC & OER (LO repository)
CESCA & MDX (LO harvester)
Interoperability issues
Use of semantics for describing LOs
Road map & future work
Conclusions
Some references
Learning objects (LO)

LOs are small pieces of content that are supposed to help learners to acquire a specific learning goal.

LO (digital) = content + metadata

Reuse, Redistribute, Revise, Remix

Important issues: Granularity / size, Open format, License

LORs: Repositories are a way to organize LOs
Table of contents

Introduction

UOC & OER (LO repository)

CESCA & MDX (LO harvester)

Interoperability issues

Use of semantics for describing LOs

Road map & future work

Conclusions

Some references
UOC & OER

UOC is an on-line distance university with more than 40,000 students and 2,500 staff.

OER (http://oer.uoc.edu) is part of a large project with the aim of promoting the development and acquisition of competences though the use of LORs.

OER is always beta, to test on research and innovation (currently now, about Statistics).
LOR goals

Repositories main goals:

Ensure **preservation**

Promote **reutilization**

Dissemination / positioning

And a LOR should become an...

**Active element of the learning process**, so students using LORs should achieve a set of competences
OER content

Exercises  PDF, QTI, ...
Examples    PDF, PPT, ODP, ...
Graphics    JPEG, PNG, ...
Simulations Applets, Flash
Data        XLS, SPSS, ...
Equations   LaTeX, MathML, ...
Tables      PDF, XLS, ...

Searching and browsing actions are designed from a teaching perspective (not librarian)
New user interface

Three complementary elements:
List of competences
Tag cloud of keywords
Visual taxonomy

Additional filters:
Resource type
Language

Open Repositories 2010, 6-9 July, Madrid, Spain
OER (http://oer.uoc.edu)
Table of contents

Introduction
UOC & OER (LO repository)
CESCA & MDX (LO harvester)
Interoperability issues
Use of semantics for describing LOs
Road map & future work
Conclusions
Some references
CESCA

CESCA is a public consortium which is integrated by the Generalitat de Catalunya, Talència, nine Catalan universities, and CSIC.

The objective is to manage e-infrastructures, in order to provide support to universities and research, based on five activity areas: communications networks; portals and repositories for university information; HPC and data-storage systems, promotion of the use and benefits of these technologies, and operation and maintenance of the entire Centre infrastructure.
CESCA & MDX

Dissertations (2000, migrating to DSpace)
Research documents (2005, DSpace)
Scientific, cultural and academic journals (2006, OJS)
Catalan digital heritage (2006, Heritrix…)
Spanish scientific and technological journals (2008, OJS)
LOs (2009, DSpace)
Documents of cultural interest (2010, DSpace)

...with/for CBUC, BC, FECYT, Catalonia Government

Open Repositories 2010, 6-9 July, Madrid, Spain
MDX (www.mdx.cat) is a cooperative repository that contains digital materials and resources resulting from teaching activities carried out in member universities. The purpose of MDX is to make the participant institutions’ teaching production more visible and widespread, thus contributing to educational innovation, on the one hand, and free access to knowledge, on the other.

Communities in MDX

- Universitat de Barcelona
- Universitat Autònoma de Barcelona
- Universitat Politécnica de Catalunya
- Universitat Pompeu Fabra
- Universitat de Girona
- Universitat de Lleida
- Universitat Rovira i Virgili
- Universitat Oberta de Catalunya
- Universitat de Vic
- Universitat Jaume I
MDX’s participant institutions

2 ways: deposit or harvest LOs

Now, 10 CBUC universities (about 1,400 LOs)
   UB, UAB, UPC, UPF, UdG, UdL, URV, **UOC**, Uvic, UJI

UOC open repositories:
O2: institutional (included)
OCW: institutional (included)
OER: thematic (not included, **yet**)

Open Repositories 2010, 6-9 July, Madrid, Spain
Table of contents

Introduction
UOC & OER (LO repository)
CESCA & MDX (LO harvester)
Interoperability issues
Use of semantics for describing LOs
Road map & future work
Conclusions
Some references
Visibility & OAI-PMH harvesters

The 4 R’s experience is better with more scope

LOR’s users are...

Harvester’s users are (potentially) ...

But, OER is not yet OAI-PMH compliant. Why?

Open Repositories 2010, 6-9 July, Madrid, Spain
Traditional perspective is library centered:

- Books, journals, works, … (mostly textual)
- Everything has a unique title, one or more authors
- Everything has a creation date, a source
- Almost everything is a PDF file

Main goal: easily finding a resource by using a minimum set of common descriptors

LO… What’s the title or the author of an exercise?
Options: hidden metadata, semantics, …
Semantic description of LOs

Incomplete metadata records (only keywords and taxonomies)

Ontologies provide additional metadata (computed, inherited, ...)

i.e. For exams, the ontology says that “title” is not mandatory and “date” means the date when the exam was done (not the document creation)

title = “Exam” + Degree + Subject + Semester
Semantic description of LOs

Ontology + rules provide:

An OAI-DC compliant metadata for any record by refactoring process

An extended schema with all the metadata available through the ListMetadataFormats of an OAI service
Pros & Cons

But... for providing an OAI-DCnq, is it necessary to fully implement the ontology? → Performance issues! Maybe a crosswalk is enough...

But, but there are more opportunities to exploit the ontology:

Ex. Useful information for filling the submit forms
Items related or “similar” for a given LO
Rules that say that, i.e., a text item has no “duration”
System architecture

Institutional LOR

- LO basic metadata
- Interaction Information

Ontology Driven Crosswalk

Ontology

View of the LO Repository Information

- OAI-PMH compliant metadata

Refactoring Process

- OAI-DC
- OAI-LOM

Open Repositories 2010, 6-9 July, Madrid, Spain
Table of contents

Introduction
UOC & OER (LO repository )
CESCA & MDX (LO harvester)
Interoperability issues
Use of semantics for describing LOs
Road map & future work
Conclusions
Some references
Road map & future work

Now, we are implementing a pilot project with:

LO’s ontology with rules:

- To generate DCnq metadata for OAI-PMH servers
- To help with related information in the submit forms

- Add-on to use with DSpace based on the Koutsomitropoulos semantic search code

More possible applications through the use of upper ontologies
Conclusions

LOR’s to preserve, reuse, disseminate... and learn

Reuse, Redistribute, Revise, Remix but first... find it

OAI-PMH harvesters increase LOR visibility

Semantics is a way to generate DCnq from a LOR and... much more

We are at the beginning, there are a lot of potential outputs to explore
References and bibliography


Open Repositories 2010, 6-9 July, Madrid, Spain
Thanks for your attention!

Contact information:

Ricard de la Vega  rdelavega@cesca.cat
Jordi Conesa  jconesac@uoc.edu
Julià Minguillón  jminguillona@uoc.edu

Open Repositories 2010, 6-9 July, Madrid, Spain