Interoperability issues between learning object repositories and metadata harvesters

Ricard de la Vega[1], Jordi Conesa[2], Julià Minguillón[2]

[1]Centre de Supercomputació de Catalunya (CESCA)

[2]Universitat Oberta de Catalunya (UOC)

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

Table of contents

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

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
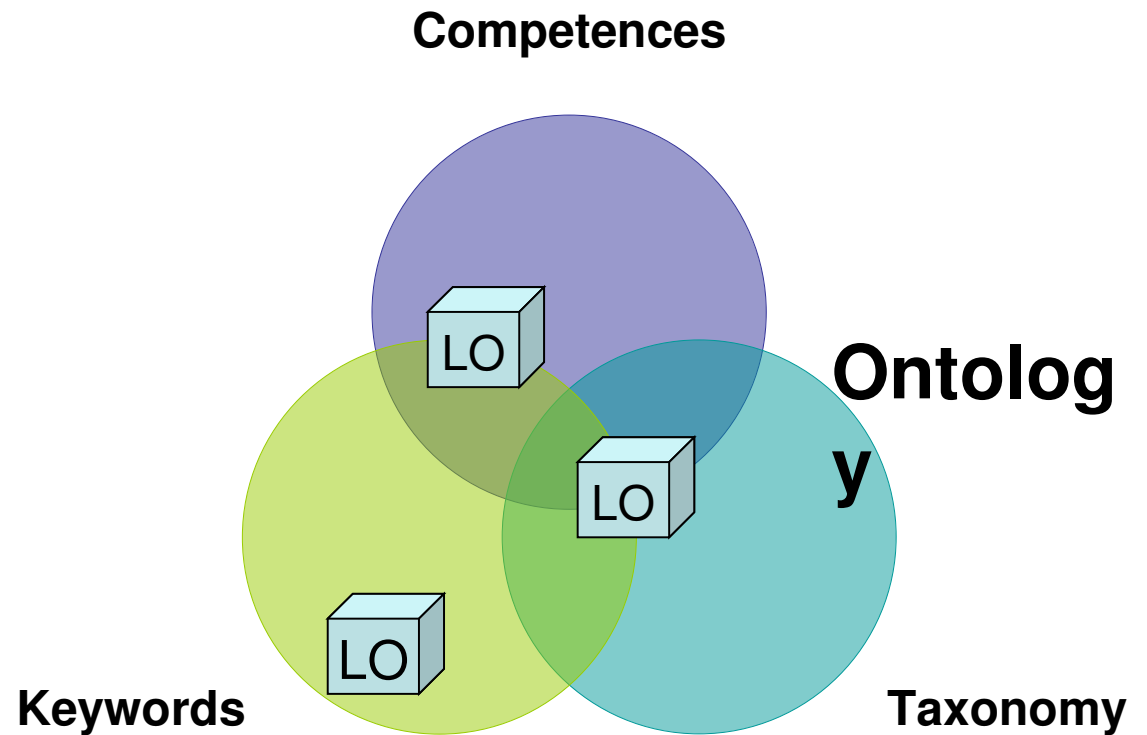
Three complementary elements:

List of competences

Tag cloud of keywords

Visual taxonomy

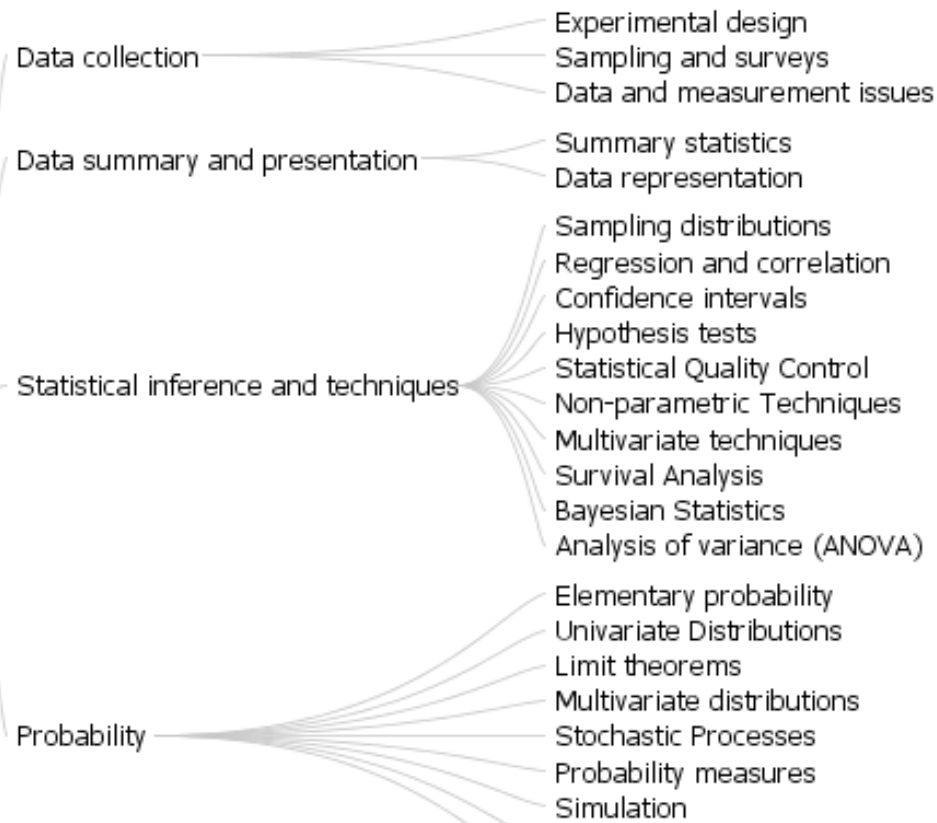Additional filters:

Resource type

Language

Competences

Ontology

Keywords

Taxonomy

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

TDR — Dissertations (2000, migrating to DSpace)

RECERCAT — Research documents (2005, DSpace)

RACO — Scientific, cultural and academic journals (2006, OJS)

padicat — Catalan digital heritage (2006, Heritrix...)

RECYT — Spanish scientific and technological journals (2008, OJS)

mdx — **LOs (2009, DSpace)**

CALAIX — Documents of cultural interest (2010, DSpace)

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

MDX (*www.mdx.cat*)

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**)

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...
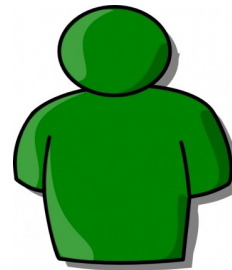
Student   Staff   Colleague
University X   University X   University Y

Harvester's users are (potentially) …

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

Traditional & OERs perspective

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$_u$

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

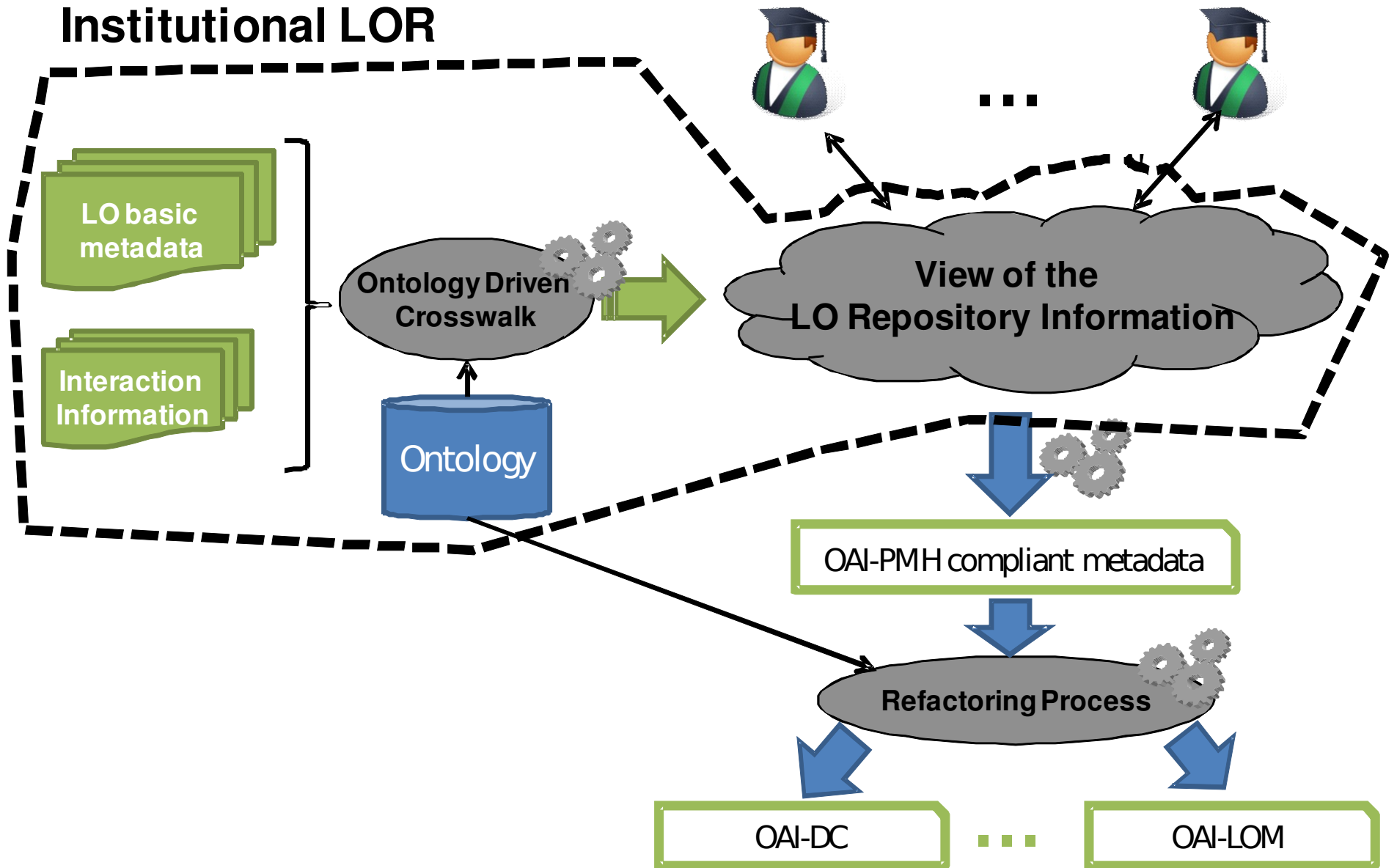Refactoring Process

OAI-DC ···· OAI-LOM

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 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

Antoniou, G., Harmelen, F. V. (2004) "A Semantic Web Primer". Massachussets Institute of Technology Press.

Córcoles, C., Minguillón, J. and Lamb, B. (2009) "Building an open social learning community around a Dspace repository on Statistics", In: Proceedings of the 4th International Conference on Open Repositories (posters), Online Version, http://hdl.handle.net/1853/28480

Currier, S. (2008), "Metadata for learning resources: an update on standards activity for 2008", ARIADNE, Vol. 55. Online Version, http://www.ariadne.ac.uk/issue55/currier

Ferran, N., Guerrero-Roldán, A., Mor, E., and Minguillón, J. (2009) "User Centered Design of a Learning Object Repository", In: Proceedings of the 1st international Conference on Human Centered Design: Held As Part of HCI international 2009 (San Diego, CA, July 19 - 24, 2009). M. Kurosu, Ed. Lecture Notes in Computer Science, vol. 5619, pp. 679-688.

Huguet, M., Anglada-i-de-Ferrer, Lluís M. and de-la-Vega, R. (2007) "Catalan Policies and Experiences on Cooperative Repositories", In: 1st Iberian Grid Infrastructure Conference Proceedings (IBERGRID). Santiago de Compostela, Spain. May 14-16, 2007. p. 63-75. Online Version, http://hdl.handle.net/2072/4083.

Kim, H.-L., Decker, S. and Breslin J. G. (2010) "Representing and sharing folksonomies with semantics", Journal of Information Science, 36 (1), pp. 57–72.

Koutsomitropoulos, D., Solomou, G., Papatheodorou, T. (2009) "Metadata and Semantics in Digital Object Collections: A Case-Study on CIDOC-CRM and Dublin Core and a Prototype Implementation", Journal of Digital Information, North America, 10, dec. 2009. Online version, http://journals.tdl.org/jodi/article/view/693.

Sicilia, M.A. and García-Barriocanal, E. (2005) "On the Convergence of Formal Ontologies and standardized e-Learning", Journal of Distance Education Technologies, 3(2), pp. 12-28.

Wiley, David A. (2000), "Connecting Learning Objects to Instructional Design Theory: A Definition, A Metaphor, and A Taxonomy", in Wiley, David A. (DOC), The Instructional Use of Learning Objects. Online Version, http://reusability.org/read/chapters/wiley.doc

Thanks for your attention!

Contact information:

Ricard de la Vega rdelavega@cesca.cat

Jordi Conesa jconesac@uoc.edu

Julià Minguillón jminguillona@uoc.edu

CC-BY-NC-SA