

OER Search

Brandon Muramatsu

mura@mit.edu

MIT, Office of Educational Innovation and Technology

December 2010

A brief bit of history...

*How can we improve the discoverability
and reuse, of OCW courses?*

*— OCW Consortium Technology Committee
2006-2007*

Our goals...

2006-2007

- **Something that was easy to implement**
 - That didn't require a lot of technical knowledge, or wading through metadata specifications
- **By using web technologies we wouldn't isolate the metadata/content**
 - This wouldn't just be for OCW Consortium...
 - It wouldn't be an "education-only" solution...
 - But it would also enable content syndication with the web-at-large
- **(There are new developments and suggestions within the OCW Consortium which we'll hear from)**

Recommendation: Expose Course Level Metadata using Dublin Core via RSS

April 2007

□ Relatively simple to implement

- OCW sites were either crazy well funded (MIT and other Hewlett Grantees) or were doing things on a shoestring (everyone else)
- RSS was (is) a simple to use/implement and common technology
- Everyone else (the rest of the Web) was doing it, this wouldn't be a specialized solution

□ Dublin Core as a metadata element set

- Dublin Core has a base level of descriptive metadata, and it could describe an OpenCourseWare course
- Dublin Core was understandable to novices
- IEEE Learning Object Metadata is fairly complex and would provide more description, but we felt the likelihood of getting lots of sites doing high quality metadata was low

Recommendation (cont.)

April 2007

▣ Course level metadata

- ▣ OCW's share courses and course materials
- ▣ Generating course level metadata might be challenging enough for existing OCWs, so we balanced getting lots of something versus getting very little, but high quality

It wasn't perfect, but was a start...

April 2007

□ We thought that...

- This provided a base level of metadata to expect
- Automatic metadata generation might help (either with course level or item level metadata, coupled with metadata inheritance)
- Other Web 2.0 technologies and techniques might help pick up where formal, descriptive metadata left off (tagging, recommenders, etc.)
- Search engines (or direct linking) were probably going to be the primary source of traffic
- These efforts would help OCWs describe materials for their own use and might encourage intra-OCW discovery mechanisms

We knew that...

April 2007

- ▣ **IEEE Learning Object Metadata would provide richer metadata**
 - ▣ Or other element sets (METS, etc.)
- ▣ **OAI-PMH, or even distributed search, would provide more robust access to metadata**
 - ▣ But it would be YAS (yet another server) to implement and run
 - ▣ It might still require a 1:1 relationship/negotiation between provider/consumer (NSDL was spending quite a bit of time and effort “curating” OAI-PMH feeds and metadata normalization requiring time and effort)

What did the metadata recommendation enable?

April 2007

- **Standardized course and item-level metadata**
 - To “share” across implementations
 - Some, like MIT and OpenLearn implemented richer metadata element sets (application profiles of IEEE LOM)
 - For many others, common metadata available via eduCommons platform
- **RSS feeds led to services**
 - OCW Finder, OER Recommender/Folksemantic.com,
 - OCWC Search (original)
 - Aggregated provider lists/feeds for CC search and OER Commons (original)

If we're going to ask folks to spend their limited time and resources, what's the most effective and impactful thing that we can ask them to do?

Philosophies and Issues

- ▣ **Metadata versus (?) content**
- ▣ **Curated vs. the open Web**
- ▣ **What's an OER? Who defines OER?**
 - ▣ Gaming the system
- ▣ **Education-specific versus Web-at-large**
- ▣ **Reduce the burden on the provider**
- ▣ **Use existing means**
 - ▣ Expressions: Creative Commons and Dublin Core
 - ▣ Mechanism: RDF, meta tags, XML, others?

My Google Wish List...

My Wish List...

- **For the last 5 years or so, I've thought that some Google engineer would come up with a way to present relevant search results to educators looking for educational materials...**
 - I don't think more or specialized repositories are the answer
 - I don't think curation is the answer
 - I don't think producer developed metadata is the answer
 - ...perhaps today is that day...

And way back when I wondered what happens if we just limit search to .edu and .k12.us (and other global equivalents, i.e., ac.uk)...and then expand with links those pages link to...

My Wish List... (cont.)

- **For the last couple years I've wondered if Google's help could be enlisted to develop/implement “learning” analytics**
 - That is more than web analytics...
 - That provides better proxies of “educational” use, such as content consumption of educational resources (say reading time coupled with path)
 - That also looks at reuse (additional occurrences of CC-licensed content on the open web, or likely derivative works)

The original question...

*“How are we [OEIT] trying to solve the
OER search problem?”*

What's OEIT working on?

- Experiments to extend the local discoverability of MIT OpenCourseWare materials
- Searching through MIT OCW Videos via their transcripts
 - Proxy for search through video (convert video to text transcript)
 - ...aka...what YouTube's Auto-Caption could be doing...
 - Integrated with MIT's Google Search Appliance, the mechanism MIT OCW uses to search their site
 - Using SpokenMedia generated transcripts for those MIT OCW videos without 99% accurate transcripts

What's OBIT working on? (cont.)

- ▣ **Local version of Folksemantic.com with “curated” collection of resources**
 - ▣ MIT OCW -> MIT OCW courses/content only
 - ▣ Could be expanded to other selected sites (via RSS feeds and/or OAI-PMH)

Thank You!

Brandon Muramatsu

mura@mit.edu

MIT, Office of Educational Innovation and Technology



Some Additional History...

How do we describe resources? (Metadata)

- ❑ MARC Records (NEEDS, through ~1994)*
- ❑ ARIADNE Metadata (through ~1997)*
MERLOT Metadata (circa 1997)*
NEEDS Metadata (through ~1997)*
- ❑ Dublin Core (activity begins 1995, v1.0 1998)
- ❑ IMS Learning Resource Metadata (adopted 1997)*
- ❑ IEEE Learning Object Metadata (standardized 2002)*
- ❑ **The world's users moves to search engines to find resources... (really sometime circa 2000)**

* Erik Duval and/or Brandon Muramatsu can speak to this

“Federating” Educational Resource Collections (not just OERs)

- **ARIADNE (1990s to present)***
 - Distributed search and union catalog
 - ARIADNE and later IEEE LOM metadata
- **National Science Digital Library (2002 to present)***
 - Union catalog
 - Primarily Dublin Core via OAI-PMH
- **GLOBE (circa 2000 to present)***
 - ARIADNE, MERLOT, SMETE.ORG, others
 - Distributed search and union catalog
 - Metadata as implemented by partners, handled by distributed search
- **Other EU Projects**

* Erik Duval and/or Brandon Muramatsu can speak to this

“Federating” Open Educational Resource Collections (OERs)

- **OCW Finder (circa 2006)***
 - Proof of concept via screen scraped MIT OCW data, then via RSS feeds
 - Transformed into OERRecommender / Folksemantic.com
- **OER Commons (2007)**
 - Lisa Petrides
- **Creative Commons (circa 2007, DiscoverEd launched 2009)**
 - Nathan Yergler
- **OCW Consortium (circa 2008)**
 - Clay Whipkey
- **OERsearch (circa 2010)**
 - Pierre Farr

* Brandon Muramatsu can speak to this