Serendipity a platform to discover and visualize Open OER Data from OpenCourseWare repositories

Edmundo Tovar (UPM, @etc91)

Nelson Piedra (@nopiedra), Janneth Chicaiza, Jorge López (UTPL)

LJUBLJANA, SLOVENIA - 23 - 25 April 2014
Theme: “Open Education for a Multicultural World”
UPM GICAC Research Group
Agenda

• The context
• Linked Open OER Data Exploitation
• Services that support the model
  – Serendipity, search of engine
  – Serendipity, maps
I. The context

• In the last years, the amount of Open Educational Resources (OER) on the Web has increased dramatically.

• The potential of this vast amount of resources is enormous but in most cases it is very difficult and cumbersome for users (teachers, students and self-learners) to visualize, explore and use this resources,

• Semantic Web technologies and, more precisely, Linked Data are changing the way information is stored and exploited, and they create bridges
II. LINKED OCW/OER DATA EXPLOITATION MODEL

Linked Data—particularly data available using open licenses—has an important role to play in information systems and could be a key feature for Open Education based on OER data on the Web of Data.
II. Linked OCW/OER Data Exploitation Model

The generation and use of linked data from heterogeneous open access sources will create new opportunities for OER and OCW initiatives. The exploitation model pipeline is organized in four stages:

1. RDF data generation:
   - The collection under study consisted of approximately 8,000 OCWs and 100,000 OERs in the collection of the OCW-Dataset of LOCWD Project.

2. Analytical approach:
   - Technically, at this stage is important define SPARQL statements to extract the data which contain the information required for each application. The applications to visualize, explore and use OCW/OER data are generated depending on the data extraction and visual configuration.

3. Visual abstraction
   - The purpose of this stage is to condense the data into a displayable way for particular application (information that is visualizable on the end-user application). Because the linked data on the Web are accessible via services or public query libraries, OCW/OER data can be processed, reused, combined, integrated, and used for several purposes.

4. Visual representation:
   - The purpose is connect LOCWD dataset with visual tools to consume and display linked OER/OCW data. This stage processes the visualization abstractions in order to obtain a visual representation of end-user applications. In this stage, the result of the process is presented to the user, e.g. faceted search, data visualizations: plot, treemap, map, timeline, etc.
The power of Linked Open Educational Resources Data

The domain of OER and OCW resources described as Linked OpenCourseWare Data holds the potential to move OER and OCW collections out of their silos and therefore opening the data:

- To leverage the knowledge capital represented by our OCW repositories
- To enrich our information landscape, to improve visibility
- To improve ease of discovery open academic resources
- To improve ease of consumption and reuse of OER & OCW
- To reduce redundancy in search of OER & OCW
- To promote innovation and added value to Open Educational Initiatives.
III. SERVICES SUPPORTED BY THE MODEL
III. Services

• Two services to explore and interact with the LOCWD data through different ways:
  – (a) **Serendipity a faceted search engine** (http://serendipity.utpl.edu.ec/), and
  – (b) **Serendipity Maps** (http://serendipity.utpl.edu.ec/map), a service of OER data visualizations.

• These applications allow users to obtain an overview of RDF datasets from OCW and OER
In Serendipity, facets correspond to properties of the OpenCourseWare content. The data are derived by analysis of the text of an item using entity extraction techniques or from pre-existing fields in a database such as author, descriptor, language, and format.

CASE 1. SERENDIPITY, A FACETED SEARCH FOR OER AND OCW RESOURCES
Context

• Search is among the most disruptive and innovative technology of the Internet. OER search is not a solved problem.

• **Facet browser interfaces provide a convenient and user-friendly way to navigate** through a wide range of Open Educational Resources Data collections:
  – To support flexible navigation through the information space
  – To refine providing suggestions of exploration, choices at each point in the search process.
Needs of stakeholders

• Users:
  – to filter content using multiple category or taxonomy terms at the same time. Users want to combine text searches, category term filtering, and other search criteria.
  – Self-learners don't know precisely what they can find on OCW site, or what to search for. Self-learners are trying to discover relationships or trends between OER & OCW.

• OCW providers:
  – To improve limitations of their own searching tools: users often get empty result sets when searching your site.
Serendipity a Faceted Search for OpenCourseWare Content

• (http://serendipity.utpl.edu.ec/)

• This service enables the faceted exploration of OER related to OCW in an OpenCourseWare search engine.

• Moving towards a Web of Linked Data, Serendipity provides a service to explore refining by facets large OCW and OER collections.
Exploring OCW in an integrated and incremental way, from any of the repositories of institutions that publish OER and OCWs.
Serendipity – expanding the original query

- **DISCOVER**: The user, when presented with the facets, is likely to discover new facets of the query that they were not aware of before.
- **REFINE**: When clicking on a facet, they will narrow down their search by expanding the original query with the suggested facet.
- **BETTER RESULTS**: more accurate and complete results, since it locates OERs using different metadata and data elements.
Accessing to a particular course. The first section contains general data about the selected course.

- On the right-hand side the traditional image search results are shown.

- On the left-hand side there is the list of facets for OCW and OER resources. In this case, a list of prominent resources in Serendipity is displayed.
Serendipity demonstrates following key features: grouping search results by facet; displaying a total number of OER & OCW per facet value, refining search results by facet value, update of the facet menu based on refined search criteria, displaying of the search criteria in a Bread Crumbs (navigation guides), ability to exclude the chosen facet from the search criteria, ability to improve ease of discovery open academic resources, ability to improve ease of consumption and reuse of OER & OCW, and ability to reduce redundancy in search of OER & OCW.

Accessing to a particular course. The second section contains information about publisher.
Accessing to a particular course. The third section provides access to OERs from the selected course.
CASE 2. SERENDITITY  DATA VISUALIZATION
Context

- [http://serendipity.utpl.edu.ec/map/](http://serendipity.utpl.edu.ec/map/)

- OVERVIEW OF DATASETS: Visualization techniques to the OER and OCW data to explore large amounts of data and interact with them: an overview of the datasets, their main types, properties and the relationships between them.

- UNFORESEEN DATA: unique opportunity on the Open OER Data environment generated by LOCWD and Serendipity.

- We generate data structures automatically from reused vocabularies or semantic representations. This will enable users to explore datasets even if the publisher of the data does not provide any exploration or visualization means.
As an important feature of Serendipity, Serendipity MAPs is a data visualization of OERs that allows users visualize data of OCW / OER / MOOC / OEP / OA Projects / OA Repositories from a dataset based on Linked Data technologies show Points of Interest, POIs).

OER Data Visualization: Map of Open Educational Repositories distributed in the world and described on LOCWD
Serendipity Maps

• Serendipity Maps use icons to represent different categories of data on a map graphically. For example, a point of interest, or POI, is an OER specific point location that someone may find useful or interesting. A description for the POI is usually included, and other information such as description, number of resources, contact information, language, license or a link to dbpedia/freebase may also be attached.
A point of interest specifies, at minimum, the latitude and longitude of the POI, assuming a certain map datum (extracted from dbpedia datasource). An example is a point on the Earth representing the location of the Massachusetts Institute of Technology, or a point on Spain representing the location of an OCW University.

An example de Serendipity POI
Geo Data Visualization of resources by tag
Serendipity - **Open data** contains valuable information that will drive insights, innovations, and discoveries, but it can be difficult to access and digest. Using data visualization, we’re simplify the complexity and drive a deeper understanding of the open educational context.
The new generation of Open Educational Resources

Linked Data vision enables a new generation of open educational resources that can be semantically described and connected with other data and discoverable sources.
Conclusions

• For effective use, reuse, discovery, distribution and sharing of OER and OCW repositories must provide a declarative query interface that supports complex expressive Web queries.

• Faceted search and data visualizations are important aspects to visualize, explore, and use Open Educational Resources Data.
Thanks!

OCWC Global Conference 2014
Theme: “Open Education for a Multicultural World”
http://conference.ocwconsortium.org/2014/

LJUBLJANA, SLOVENIA - 23 - 25 April 2014
UPM GICAC Research Group

Edmundo Tovar (UPM, @etc91) - Nelson Piedra (UTPL, @nopiedra)