A LOD SEARCH INDEX STRUCTURE FOR LEARNING RECOMMENDATION RESOURCES SYSTEM

Michelle Wang
Research Center for Information Technology Innovation, Academia Sinica

2013/5/29, Hong Kong
OUTLINE

- Learning Resources Recommendation System
  - LOD: format, example, application
- Open Data Business / Ecosystem
- Combine Resources Recommendation System and LOD
Learning Resources Recommendation System

- Learning System Trends:
  - Personalization
  - Course Adaptive
  - Knowledge Map
  - Mobility
Learning Resources Recommendation System

- Main challenge:
  - data acquisition problem
  - knowledge acquisition problem
  - open corpus problem
  - scalable
  - adaptive
  1. related data from different domains
  2. identical users in different communities
Learning Resources Recommendation System

- Tim Berners-Lee's 5 star Open Data plan:

  - ★★★★★ make your stuff available on the Web (whatever format) under an open license
  - ★★★ make it available as structured data (e.g., Excel instead of image scan of a table)
  - ★★★★ use non-proprietary formats (e.g., CSV instead of Excel)
  - ★★★★★ use URIs to denote things, so that people can point at your stuff
  - ★★★★★★ link your data to other data to provide context
OPEN DATA - format

In general, RDFa uses simple attributes in XHTML tags (often <span> or <div>) to assign brief and descriptive names to entities and properties. Here’s an example of a short HTML block showing basic contact information for Bob Smith.

```html
<div>
    My name is Bob Smith but people call me Smithy. Here is my home page:
    I live in Albuquerque, NM and work as an engineer at ACME Corp.
</div>
```

Here is the same HTML marked up with RDFa.

```html
<div xmlns:v="http://rdf.data-vocabulary.org/#" typeof="v:Person">
    My name is <span property="v:name">Bob Smith</span>,
    but people call me <span property="v:nickname">Smithy</span>.
    Here is my homepage:
    I live in Albuquerque, NM and work as an <span property="v:title">engineer</span>
    at <span property="v:affiliation">ACME Corp</span>.
</div>
```

http://support.google.com/webmasters/bin/answer.py?hl=en&answer=146898
一般而言，RDFa 会使用 XHTML 标记中的简单属性（通常是 `<span>` 或 `<div>`）为实体和属性指派简要的描述性名称。以下是一个简单的 HTML 程序段示例，内容是 Bob Smith 的基本联络资讯。

```html
<div>
    我是 Bob Smith，不过朋友都叫我 Smithy，以下是我的首页网址：
    <a href="http://www.example.com">www.example.com</a>。
    我住在 Albuquerque (NM)，目前服务於 ACME Corp.，担任工程师一职。
</div>
```

以下是使用 RDFa 为相同的 HTML 内容加上标记后的程式码：

```html
<div xmlns:v="http://rdf.data-vocabulary.org/#" typeof="v:Person">
    我是 <span property="v:name">Bob Smith</span>，
    不过朋友都叫我 <span property="v:nickname">Smithy</span>。
    以下是我的首页网址：
    <a href="http://www.example.com" rel="v:url">www.example.com</a>。
    我住在 Albuquerque (NM)，目前服务於 <span property="v:affiliation">ACME Corp</span>，
    担任 <span property="v:title">工程师</span>一职。
</div>
```
OPEN DATA - format

Source: Applying Linked Open Data for Green Introduction
OPEN DATA - example

http://data.gov.uk
OPEN DATA - example

http://data.gov.tw
OPEN DATA - example

http://open-data.europa.eu/
OPEN DATA - application

- World Bank DataFinder Mobile Apps

- real estate listings
  - http://www.zillow.com/

- 等公車
  - App i84 Taipei 大台北公車
  - 台北等公車
OPEN DATA - application

Open Data Business / Ecosystem

Why LINKED Open Data???

Cost of using data sources

Number of data sources

Traditional data integration process

Ontology-driven process

Source: Price Waterhouse Coopers – Technology Forecast, Spring 2009
Open Data Business / Ecosystem

- OGD BusinessDay 2012: Open Data (&) Business, 22.03. 2012, Vienna
  - http://ogdb.eventbrite.com/

- European Data Forum 2012, 6-7 June12, Copenhagen, DK
  - http://data-forum.eu

- OGD2012 Conference, Linz, Austria, 26.06. 2012
  - http://www.ogd2012.at


- OKFO / OGD Austria (DE):
  - http://opendata.at/okfo

- Health Data Initiative
  - http://www.hhs.gov/open/initiatives/hdi/

- Open growth: Stimulating demand for open data in the UK

- Open Data Center Alliance
  - http://www.opendatacenteralliance.org/newsroom/mediaresources
Open Data Business / Ecosystem

Source:
http://www.hhs.gov/open/plan/opengovernmentplan/initiatives/initiative.html
http://radar.oreilly.com/2010/06/unlocking-innovation-through-d.html
Open Data Business / Ecosystem

Combine Resources Recommendation System and LOD

- **Challenges**
  - Data level integration and connection to Linked Open Data cloud.
Combine Resources Recommendation System and LOD

Linked Services (mEducator3.0)
## Combine Resources Recommendation System and LOD

### Open Data Business - Sectors

<table>
<thead>
<tr>
<th></th>
<th>Public Sector</th>
<th>NPO &amp; NGO</th>
<th>Private Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Integration (BI, MI, Data Warehouse)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Example: Industrial Facilities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Example: Media &amp; Publishing</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Example: Real Estate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Example: Transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Applications (APPS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Direct revenues through license fees for data re-use</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Direct tax revenue through new job creation, better economic development</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Direct revenues for the economy on top of better &amp; innovative products &amp; services</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Enrichment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sale of enriched data on top of open data</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enrichment of data by meta data (increase in value of data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Converter services for harmonised data formats (ETL, Data Integration Services)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Visualisation &amp; -analysis Services</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Services in the fields of data analysis and data visualisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Infrastructure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cloud services</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLAs for enterprise data management</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Open Data Innovation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open Innovation via crowd sourcing mechanisms</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open Innovation via mass customisation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Combine Resources Recommendation System and LOD

- **Index Integrated**
  - Search Engine, ex: Google, knowledge map
  - Application, ex: DBPEDIA
  - Search tool, ex: Solr

- **Features for Recommendation System**
  - user preferences
  - Item / resource content
  - Usage / use log
  - Interactions / social

- **Learning Course / Resources at LOD**
  - Course -> Knowledge -> semantic -> user model
  - personal knowledge recommendation
  - Linked Course
Combine Resources Recommendation System and LOD

- Challenges of LOD measurement
  - reliability
  - peak-load
  - performance
  - usefulness
  - attacks
THE LIFE-CYCLE OF LINKED DATA

- Extraction of RDF from text, XML and SQL
- Querying and Exploration using SPARQL
- Authoring of Linked Data using a Semantic Wiki
- Semi-automatic link discovery between Linked Data sources
- Knowledge-base Enrichment and Repair
Combine Resources Recommendation System and LOD

- **DBPEDIA**
  - Cinemappy

Source: Cinemappy: a Context-aware Mobile App for Movie Recommendations boosted by DBpedia
Combine Resources Recommendation System and LOD

Social filtering + Collaborative Filtering (user-based + content-based)
Combine Resources Recommendation System and LOD

Architecture of Linked Course Data

Data Conversion Layer
- extract the concept, property, instance and relation among concepts in RDF triples.
- Output: RDF data

Data Enrich and Linked Layer
- map the data to the corresponding entities in linked data
- output: discover the potential relationship between entities

Data Storage and Indexing Layer
- unify the data formats in order to avoiding the error caused by the different formats.

Data Application Layer
- show the linked data with the text, tables and image from different views according to users needs.

Source: Linked Course Data-based User Personal Knowledge Recommendation Engine
Conclusions

- Optimize data for search engine and application
  - Transfer resources to RDF / LOD format, combine with existing application

- Make resources courseelize
  - Ready recommendation-related info:
    - Content
    - Author
    - Usage
Discuss

Q & A