CRITERIA FOR THE SELECTION OF A SEMANTIC REPOSITORY FOR MANAGING SKOS DATA

CONVEGNO AIB CILW 2016

Ricardo Eito-Brun
Universidad Carlos III de Madrid
MOVING TO SKOS...

- Two sides of Linked Open Data projects:
  - Moving online, convert to RDF, our data.
  - Give other people the possibility of using your data (we mean “other software applications”).

- Today, most organizations focus on data conversion and publishing
  - Convert Thesauri, Classification schemas or KOS to SKOS.
  - The results of this conversion are published..., as HTML data..., not reusable by machines.
  - Set up SPARQL end-points, that are hard to use
    - language syntax and
    - Complex, technical integration of “consumer applications” with SPARQL end points in specific context (e.g. non-Java desktop apps)
  - The last two points are done with the help of RDF repository/DB tools.
TOOLS FOR MANAGING SW-DATA

- Functions related to data load / entry:
  - Bulk Import of RDF data.

- Functions related to data extraction / search
  - SPARQL end-points.
  - Full text indexing
  - Proprietary APIs

- Main problems:
  - Lack of Integration with the tools we use for editing RDF/SKOS data.
  - Data access restricted to complex SPARQL
  - Not all the data we must manage are RDF-based (e.g. EAC authority records)

- This scenario led to the selection of an XML database tool for managing SKOS data.
INTEGRATION OF SKOS AND SRU IN A DISTRIBUTED COLLABORATION ENVIRONMENT FOR EAD DESCRIPTION

1. Wider visibility to existing thesauri & classification systems
2. Leverage investments in the development of indexing languages.
3. Open, standard-based access to indexing languages.
4. Integration and reuse from metadata creation tools is easier.
5. New semantic integration / distributed search capabilities between repositories based on controlled vocabularies.

Cataloguer proceeds to assigning descriptor
Cataloguer selects SKOS DB and search criteria.
SRU search requests is directed to remote web server.
Response data are shown to end-user, who can assign descriptor and see details.

Request is processed and response sent back to metadata editor.

SKOS files are generated by a conversion and uploaded in an Oracle XML DB database.

XML Metadata Editor

http://www.uc3m.es/bibdoc/sruSrvr/skos/processRequest.php?
version=1.2
&operation=searchRetrieve
&query=CQL SEARCH
&maximumRecords=100
&recordSchema=skos_summary

Web Server

SKOS DB

1. Wider visibility to existing thesauri & classification systems
2. Leverage investments in the development of indexing languages.
3. Open, standard-based access to indexing languages.
4. Integration and reuse from metadata creation tools is easier.
5. New semantic integration / distributed search capabilities between repositories based on controlled vocabularies.

Ricardo Eito-Brun (UC3M)