Common Access to EU Information based on semantic technology

Presented by: Peter Schmitz
Date: 15/03/2012
Who is the Publications Office of the EU?

Common access to EU information based on semantic technology - CELLAR Project

- Vision and target architecture
- Use of semantic technology
- Multilingual aspects
- Interoperability
Who is the Publications Office of the EU?

In summary: From the EU to the public
Who is the Publications Office of the EU?

What do we publish?

- “Official” publications
- Other “general” publications
Who is the Publications Office of the EU?

Key production indicators for 2011

- **Official Journal of the European Union**
  - 866 issues, 22/23 languages, five days a week, 1,065,425 pages
  - 411,850 public procurement notices
  - Consolidation of EU Law (52,097 legislative acts)

- **General publications**
  - 5,972 titles, official languages of the EU, but also other languages

Dissemination

- Paper copies
- Online services

Long-term Preservation

- Paper archive
- Digital archive
Who is the Publications Office of the EU?

Main public online services

- **EUR-Lex**
  - [eur-lex.europa.eu](http://eur-lex.europa.eu)
  - EU Law

- **EU Bookshop**
  - [bookshop.europa.eu](http://bookshop.europa.eu)
  - General publications

- **Ted**
  - [ted.europa.eu](http://ted.europa.eu)
  - Tenders electronic daily

- **Simap**
  - [simap.europa.eu](http://simap.europa.eu)
  - eProcurement

- **CORDIS**
  - [cordis.europa.eu](http://cordis.europa.eu)
  - Community Research and Development Information Service
  - Research and development
To make available at a single place all metadata and digital content managed by the Publications Office in a harmonised and standardised way in order:

- To guarantee to the citizen a better access to law and publications of the European Union;
- To encourage and facilitate reuse of content and metadata by professionals and experts;
- To preserve content and metadata and access to contents and metadata over time.
CELLAR Project: Target architecture
CELLAR project: Use of semantic technology

Common data model: Ontology based on FRBR model

FRBR = Functional requirements for bibliographical records
Metadata represented in RDF
Taxonomies represented in SKOS
Example: Multilinguism enabled by coded metadata

 CELLAR project: Multilingual aspects

Example: Multilinguism enabled by coded metadata
CELLAR project: Interoperability

Adoption of standards

- **METS** *(Metadata encoding and transmission standard)* -> ingestion protocol
- Dublin Core -> core metadata definition
- FRBR compliant -> data model/ontology
- Linked Open Data (LOD) -> access/reuse
  - Web-friendly ("RESTful") Interface
  - RDF //owl
- Standard Query Language (SPARQL) -> access/reuse
CELLAR project: Interoperability

Definition of standards

- **Core metadata** (Restricted shared set of metadata for each resource based on Dublin Core) -> enable global search
- **Common authority tables** -> harmonize metadata
- **Exchange protocol for EU legislative procedures** -> interoperability

Under preparation:

- **European Legislative Identifier (ELI)** (initiative of EU member countries, supported by the Publications Office) -> interoperability
Thank you for your attention!