Using Linked Data as a framework for enterprise KOS

Dave Clarke
CEO Synaptica
@Synaptica
@DavidClarkeBlog

Dublin Core
Sep. 10-13, 2018
Porto, Portugal
#dcmi18
1. What is LOD-KOS
2. Examples of open data LOD-KOS
3. What is LED
4. How are enterprises using LED and LOD
5. How do enterprises manage confidentiality
What is LOD-KOS

- **LOD**: Linked Open Data – open data structured as RDF triples (subject-predicate-object) in which entities and predicates are uniquely identified by HTTP URIs
- **KOS**: Knowledge Organisation Systems – conceptual and named entity terminologies organised within a structured scheme using hierarchical, associative and equivalency relationships
- **Semantic framework**: family of W3C specifications for KOS and semantic resource description including RDF, SKOS, SKOS-XL, OWL, Linked Data Platform
- **Value vocabularies**: descriptive terminologies of concepts and named entities, e.g. lexicons, thesauri, taxonomies, classification schemes and name authorities
- **Property vocabularies**: sets of metadata elements or property types including the data properties and object properties (i.e. predicates) of an ontology
- **Lightweight ontologies**: KOS that use selections of ontological classes and properties derived from one or more property vocabulary to fit the needs of a specific community (as contrasted with reference ontologies containing semantic and axiomatic rules)
Examples of LOD-KOS

- **Individual Vocabulary Providers**: e.g. EuroVoc multilingual thesaurus of the EU [http://eurovoc.europa.eu/](http://eurovoc.europa.eu/)

- **Individual Institution Providers**: e.g. Library of Congress Data Services comprising LC Subject Headings, LC Classification, LC Name Authority File [http://id.loc.gov/](http://id.loc.gov/), or Getty Vocabularies LOD comprising the Art & Architecture Thesaurus (AAT), the Getty Thesaurus of Geographic Names (TGN), and the Union List of Artist Names (ULAN) [http://vocab.getty.edu/](http://vocab.getty.edu/)


- **Registries**: e.g. BARTOC (Basel Register of Thesauri, Ontologies & Classifications) over 2740 KOS, 315 of these are available in RDF format [https://bartoc.org/](https://bartoc.org/) and LOV (Linked Open Vocabularies) over 600 registered vocabularies; many of the vocabularies are property vocabularies [http://lov.okfn.org/dataset/lov](http://lov.okfn.org/dataset/lov)
Linked Open Data cloud

big and growing

all you need is L♥D
What about using Linked Data as a framework for enterprise KOS?

- LOD (Linked Open Data): shared openly
- LED (Linked Enterprise Data): behind the firewall
Use of Standards

- 39% use CV standards like ISO 25964
- 21% use SKOS
- 3% use SKOS-XL
- 5% use OWL
- 24% use Linked Data Platform
- 36% use metadata like Dublin Core
- 12% use Schema.Org
- 12% use LOD as reference resources

Public v. Private

- 52% behind firewall
- 36% partially exposed
- 12% made available as LOD

Expectations

- 57% Linked Data impacting their business
- 51% Semantic Web impacting their business

Survey open from pinned Tweet at: https://twitter.com/DavidClarkeBlog
How public or private are your taxonomies

- Made available for reuse: 4
- Public facing, e.g. in websites: 13
- Shared with trusted partners: 7
- Only behind firewall: 18
1. Flexible compartmentalization
2. Lightweight ontologies
3. Mix-and-match property vocabularies
4. Ease-of-use value vocabulary editing
5. Data portability
Interpretive Model of KOS

**Predicates:**
- **links** connecting a thing to another thing (object properties), or a thing to a string (data properties)

**Classes:**
- **types** of individuals that are grouped together because they share a common set of properties

**Schemes:**
- **sets** of concepts or named entities that belong together because they describe a domain of knowledge

**Concepts:**
- **individuals** (concepts or names) defined by their classes and properties and linked together within a scheme

**Ontology**
- Coherent sets of predicates and classes defining the logical structure of a domain

**Taxonomy**
- Coherent sets of terminology defining the entities within a domain

Knowledge Organization Systems
1. Manage a library of proprietary and/or public domain ontology classes and predicates

2. Design KOS schemes using plug-and-play ontology classes and predicates

3. Create and curate descriptive taxonomies that conform with the business rules of each scheme
Managing a Library of Predicates

The Simple Knowledge Organization System provides some basic vocabulary for associating lexical labels with resources of any type. In particular, SKOS enables a distinction to be made between the preferred, alternative and hidden lexical labels for any given resource.
Linked Data for Enterprise KOS

Classes

Defining Classes
Linked Data for Enterprise KOS

Schemes

Designing KOS Using Predicates and Classes
Enterprises have identifiable needs to filter Linked Enterprise Data for restricted access through both UI and also APIs and Endpoints:

1. By scheme

2. By collection

3. By individual term

4. By individual data property

5. By individual object property (relationship)

6. By approvals, languages, users

By all of the above
Ongoing research

1. What are the advantages of using the Linked Data model?
2. What are the challenges of using the Linked Data model?
3. How public facing are enterprise KOS?
4. How do enterprises approach compartmentalization?
Using Linked Data as a framework for enterprise KOS

Thank You

Dave Clarke
CEO Synaptica
@Synaptica
@DavidClarkeBlog

Dublin Core
Sep. 10-13, 2018
Porto, Portugal
@DublinCore