On practical aspects of enhancing semantic interoperability using SKOS and KOS alignment

Antoine ISAAC

KRR lab, Vrije Universiteit Amsterdam
National Library of the Netherlands

ISKO UK Meeting, July 21, London
Enhancing semantic interoperability using SKOS

Agenda

• (Optional) Semantic Web refresher
• Representing KOSs using SKOS
  • Main features
  • Practical issues
• Demo
• SKOS and semantic alignment of KOSs
The Semantic Web: a web of resources

• Pointing at resources
  
  • What? Knowledge objects
    • everything that we may want to refer to
    • including documents, persons...

  • How? Uniform Resource Identifiers (URIs)
    • E.g., HTTP URLs: http://www.few.vu.nl/~aisaac/
Enhancing semantic interoperability using SKOS

A Web of resources

myVoc: Amsterdam

theirVoc: Article

http://ex.org/files/file1

Here: URIs (namespace and localname)
Note: different locations!
Describing Semantic Web resources: RDF

- Pointing at resources: URIs
- Creating structured assertions involving resources

  - *What?* **Typed links** between resources

  - *How?* **RDF (Resource Description Framework)**
    - Statements
    - subject-predicate-object
Enhancing semantic interoperability using SKOS

Data in an RDF "graph"

theirVoc: Article

http://ex.org/files/file1

theirVoc: subject

myVoc: Amsterdam
The Semantic Web approach: A Web of (meta)data

The Netherlands hasCapital Amsterdam

Amsterdam defines type City

type

Article subClassOf Document

partOf

file1

defines paragraph3
What's the role of SKOS in this?

- Porting KOSs to the Semantic Web
  - Reminder: SKOS is for publication and access, not replacement
Enhancing semantic interoperability using SKOS

**Agenda**

- Semantic Web refresher
- **Representing KOSs using SKOS**
  - Main features
  - Practical issues
- Demo
- SKOS and semantic alignment of KOSs
SKOS (Simple Knowledge Organization System)

- SKOS offers a vocabulary to create RDF data representing KOS content
  - **Concepts** and **ConceptSchemes**
  - Lexical properties (**prefLabel**, **altLabel**)
  - Semantic relations (**broader**, **related**)
  - Notes (**scopeNote**, **definition**)

Enhancing semantic interoperability using SKOS
Conceptual resources with URIs

prefix skos: <http://www.w3.org/2004/02/skos/core#>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix ex: <http://example.org/>
Enhancing semantic interoperability using SKOS

**Labels as strings**

- USE/UF functions, as in ISO2788
- **But a concept-oriented model!**
  - Concepts are first-order (RDF) entities
  - Labels are RDF literals (simple string values)
  - Labels are linked via the concept resource
Enhancing semantic interoperability using SKOS

(Multilingual) labels

- Multilingual functions, as in ISO5964
- But a concept-oriented model, again
Enhancing semantic interoperability using SKOS

Semantic relations: broader, narrower and related

- Same function as BT/NT/RT links
- Similar intended meanings
  - e.g. **broader** can cover partitive, generic, or class-instance relationships
Enhancing semantic interoperability using SKOS

Documenting concepts

```
skos:note
|  +-- skos:definition
|     +-- skos:scopeNote
|     |  +-- skos:example
|     |     +-- skos:historyNote
|     +-- skos:editorialNote
|     +-- skos:changeNote
```
Enhancing semantic interoperability using SKOS

Example: thesaurus

animals
  NT cats

cats
  UF domestic cats
  RT wildcats
  BT animals
  SN used only for domestic cats

domestic cats
  USE cats

wildcats
Enhancing semantic interoperability using SKOS

Example: SKOS graph

- animals
  - NT cats
- cats
  - UF domestic cats
  - RT wildcats
  - BT animals
  - SN used only for domestic cats
- domestic cats
  - USE cats
- wildcats

SKOS relationships:
- skos:broader
- skos:narrower
- skos:related
- skos:prefLabel
- skos:altLabel
- skos:scopeNote
Example: RDF XML serialization

```
<rdf:RDF>
  <skos:Concept rdf:about="http://example.org/animals">
    <skos:prefLabel xml:lang="en">animals</skos:prefLabel>
    <skos:narrower rdf:resource="http://example.org/cats"/>
  </skos:Concept>
  <skos:Concept rdf:about="http://example.org/cats">
    <skos:prefLabel xml:lang="en">cats</skos:prefLabel>
    <skos:altLabel xml:lang="en">domestic cats</skos:altLabel>
    <skos:scopeNote>used only for domestic cats</skos:scopeNote>
    <skos:broader rdf:resource="http://example.org/animals"/>
    <skos:related rdf:resource="http://example.org/wildcats"/>
  </skos:Concept>
  <skos:Concept rdf:about="http://example.org/wildcats">
    <skos:prefLabel xml:lang="en">wildcats</skos:prefLabel>
  </skos:Concept>
</rdf:RDF>
```
Enhancing semantic interoperability using SKOS

Agenda

• (Optional) Semantic Web refresher
• Representing KOSs using SKOS
  • Main features
  • Practical issues
• Demo
• SKOS and semantic alignment of KOSs
How does SKOS stand the test of application?

• How much interoperability does porting to SKOS really allow?
• Are there hidden caveats?
  • Different ways to convert similar things
  • Different interpretations of SKOS constructs?
• Things impossible to convert
Issue 1: where can URIs come from?

- Generated from pre-existing identifiers
  - sh96011203
- Using (less stable) labels is not always a good idea
  - Logotypes (Printing)
**Issue 1: "web-enabled" names?**

*Do concept names have to refer to accessible documents?*

- 1. (first approach) Whatever qualifies as URI, even it does not “exist” on the web
  - http://example.org/animals
- 2. (better) Basic web-enabled name, resolves into some document
  - http://catalogue.bnf.fr/ark:/12148/cb11931683w
- 3. (ideal) URI with content negotiation to serve SKOS/RDF, HTML...
  - http://lcsh.info/sh96011203#concept
  - Cf. Best Practice Recipes for Publishing RDF Vocabularies
    - http://www.w3.org/TR/swbp-vocab-pub/
Issue 2: different types of labels

- Notice: SKOS does not offer guidelines for good labels
- But it assumes some characteristics for the different kinds of labels, that could influence conversion
  - (Hard) A concept has only one prefLabel per language
  - (Soft) No two concepts from a same concept scheme should have the same prefLabel in a given language
- Cf. notion of “descriptor”
Issue 2: classifications, notations and labels

• Notations
  - 21.51 "technique and materials"

• Can we use notations as SKOS preferred labels?
  • They (are supposed to) make sense for users & be displayed
  • They are unambiguous
**Issue 2: classifications, notations and labels**

- Captions could also be considered as preferred labels
  - They are often displayed
    - They can be ambiguous
      - But the `prefLabel` uniqueness constraint was soft!
      - Yet experts could choose to have all captions as `altLabels`...

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.00</td>
<td>&quot;painting: general&quot;</td>
</tr>
<tr>
<td>- 21.01</td>
<td>&quot;technique and materials&quot;</td>
</tr>
<tr>
<td>21.50</td>
<td>&quot;sculpture: general&quot;</td>
</tr>
<tr>
<td>- 21.51</td>
<td>&quot;technique and materials&quot;</td>
</tr>
</tbody>
</table>

Interop. at risk!
Enhancing semantic interoperability using SKOS

Semantic relations: *broader, narrower and related*

- ex:animals
- skos:broad
- ex:birds
- skos:narrow
- ex:ornithology
- skos:related
Enhancing semantic interoperability using SKOS

**Issue 3: semantics for semantic relations**

- **Is `broader` "transitive"?**
  - ex:cats → skos:broader → ex:mammals → skos:broader → ex:animals

- **If yes, we can miss the “original” information**
  - ex:cats → skos:broader → ex:mammals → skos:broader → ex:animals
Issue 3: semantics for semantic relations

- **broader** is not transitive in general
- It has a super-property **broaderTransitive** with semantics of “has ancestors”
  - 1: every **broader** statement ("parent") logically implies a **broaderTransitive** one ("ancestor")
  - 2: **broaderTransitive** is transitive!
**Issue 3: semantics for semantic relations**

- **broader** and **narrower** are inverse of each other

```plaintext
X \(\text{skos:narrower}\) Y
\(\text{skos:broader}\) Y \(\text{skos:narrower}\) X
```

- **related** is symmetric!

```plaintext
X \(\text{skos:related}\) Y
\(\text{skos:related}\) Y \(\text{skos:related}\) X
```

- Assumption: there are not many exceptions in KOSs
- A non-symmetric specialization of **related** can be coined if needed
Semantics of SKOS

- SKOS semantics make assumptions that distinguish what could be regularly inferred from a statement
  - *broader and narrower are inverse of each other*
  - *broader is transitive*

- This answers some questions about what should be explicit or not in a SKOS conversion, and what can (shall) be inferred from it
  - Important for specifying application, e.g. services
  - *Crucial for interoperability!*

- Beware: this assumes reasoning, or a simulation of it!
Example of custom extension for SKOS

- Creating a non symmetric specialization of related?
  \( \text{my:nonSymmetricRelated} \) \( \text{rdfs:subPropertyOf} \) \( \text{skos:related} \)

- Assertions of \( \text{my:nonsymmetricProperty} \) do not imply inference of reciprocal statements

- If RDFS semantics are applied (e.g. by a reasoner) there is inference of standard SKOS \( \text{skos:related} \) statements
Other KOS features which could harm interoperability

Very difficult to represent in SKOS
• Synthesis of new subjects
  • Using subdivisions: Brass bands--Sponsorship
• Links to compound non-preferred terms
  • Cf. Stella/Leonard/Nicholas

Can be represented, but not really standard
• Qualifiers in labels: Technique (painting)

Standard, but may not be used
• Groupings by Collection, cf. Doug/Ceri

• Cf. next presentations!
Enhancing semantic interoperability using SKOS

Is that damn thing useful?

- At least it's there!
  - There is a proposed standard to represent KOS on the Semantic Web

- It allows to publish KOSs
  - LCSH, Agrovoc...

- It allows to develop applications with re-usable & interoperable components
  - Cf. Doug/Ceri and Bernard
Benefit of SKOS

- Homogeneous SW representation of vocabularies (and metadata)
Is that damn thing useful?

- For most aspects of a KOS, conversion is relatively smooth

- It makes some commitments more explicit
  - Nothing compared to representation as a formal ontology
  - *Believe me!*

- A basis for (your!) experience sharing
  - Comparing conversion strategies
  - Realizing the interoperability issues there
  - Devising agreed extensions
Enhancing semantic interoperability using SKOS

Demo!

- KB Illuminated Manuscripts
- BNF Mandragore Manuscripts

Enhancing semantic interoperability using SKOS

Subject vocabulary, collection 1

Subjects

Collection

- Bibliothèque nationale de France Mandragore Collection (2770)
- Illuminated Manuscripts Collection of the National Library of the Netherlands

IconClass(Eng)

- LOCAL (2310)
- Abstract Ideas and Concepts (55)
- Bible (1819)
- Classical Mythology and Ancient History (393)
- History (200)
- Human Being, Man in General (594)
- Literature (58)
- Nature (554)
- Religion and Magic (1679)
- Society, Civilization, Culture (1952)
Enhancing semantic interoperability using SKOS

Demo (2)

Hierarchical path from root to selected subject

Possible specialisation for selected subject
Enhancing semantic interoperability using SKOS

Demo (3)

Semantic alignment of subjects activated

Collection

IconClass(Eng): Nature > earth, world as celestial
body > animals > amphibians

Found 11 objects

Document from Collection 2

These terms define your current search. Click the 'x' to remove a term.

version 0.20 (serql=sesame)

MAPPING: mandrabase mandraNewNONE

Refine your search further within these categories:

Collection

Bibliotheque nationale
Illuminated
de France Mandragore
Manuscripts Collection
Collection (5)

IconClass(Eng): all > Nature > earth, world as celestial
body > animals > amphibians

LOCAL (5) tailless amphibians (3)
**Demo (4)**

<table>
<thead>
<tr>
<th>Version 0.20 (serql=sesame)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title:</strong>  Fable : le cobra et les grenouilles</td>
</tr>
<tr>
<td><strong>Picture:</strong> <img src="http://visualiseur.bnf.fr/Visualiseur?Destination=Mandragore" alt="Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Image taille réelle:</strong></th>
<th><a href="http://visualiseur.bnf.fr/Visualiseur?Destination=Mandragore">http://visualiseur.bnf.fr/Visualiseur?Destination=Mandragore</a></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MANDRAGORE:</strong></td>
<td>plante</td>
</tr>
<tr>
<td><strong>MANDRAGORE:</strong></td>
<td>naja</td>
</tr>
<tr>
<td><strong>MANDRAGORE:</strong></td>
<td>grenouille</td>
</tr>
<tr>
<td><strong>MANDRAGORE:</strong></td>
<td>fable</td>
</tr>
<tr>
<td><strong>MANDRAGORE:</strong></td>
<td>arbre</td>
</tr>
<tr>
<td><strong>DATE:</strong></td>
<td>13e siècle</td>
</tr>
</tbody>
</table>

Subject from voc2 aligned to voc1:amphibians”
Enhancing semantic interoperability using SKOS

Demo: noticeable facts

- KOS-independent interface
  - The French vocabulary has just “replaced” an English vocabulary that was used in a previous pilot

- Makes use of standard SKOS constructs
  - `broader`, `prefLabel`

- Can exploit standard alignment relations

- Semantic equivalence can be computed thanks to SKOS' seamless representation of multilingual labels
  - It’s actually a case of French-to-French alignment!
Enhancing semantic interoperability using SKOS

Agenda

- Semantic Web refresher
- Representing KOSs using SKOS
  - Main features
  - Practical issues
- Demo
- SKOS and semantic alignment of KOSs (time?)

No time?
Aligning vocabularies
Vocabulary Alignment

• Aim: find correspondences between different concepts with comparable meanings

• Doing it manually or (at least semi-)automatically

• Cf. ontology alignment in Semantic Web research
  • Lexical
  • Structural
  • Statistical
  • Background knowledge still is a difficult research problem!
Enhancing semantic interoperability using SKOS

Mapping concepts with SKOS

- ex1:animal
- ex1:platypus
- ex2:animals
- ex2:eggLaying Animals
- ex2:eggs

- skos:exactMatch
- skos:broadMatch
- skos:relatedMatch
SKOS contribution to mapping

- A common way to represent important info for KOS use cases
  - Focusing on types of mapping relationships
- Note: can be used in combination with more complex formats developed by the ontology alignment community
  - E.g. to give mappings a confidence measure

- Again with (debatable?) semantics
  - `broadMatch` is a sub-property of `broader`
  - Allows to seamlessly use mappings as basic KOS relationships
  - Still keeps the difference at the statement level

```
ex1:platypus has skos:broadMatch to ex2:eggLaying Animals
```

```
ex1:platypus has skos:broader to ex2:eggLaying Animals
```
Conclusion

- Representing KOSs using SKOS
  - Main features
  - Practical issues
- Demo
- SKOS and semantic alignment of KOSs

→ Despite some issues, SKOS provides a crucial contribution to enhance interoperability of KOSs
Thank you!

- SKOS site: http://www.w3.org/2004/02/skos
- Reminder: Comments highly welcome on
  - SKOS Reference: http://www.w3.org/TR/skos-reference
  - SKOS Primer: http://www.w3.org/TR/skos-primer