Interoperability of KOS Metadata Schemas in BARTOC and JSKOS

Jakob Voß  Andreas Ledl

16th European NKOS Workshop, Copenhagen, 2016-10-16
Interoperability of KOS metadata schemas (2016-10-16)
Founding Reasons

- Make traditional resources more visible
- Provide key characteristics of resources
- Encourage human and machine assessment of these resources by applicability to semantic projects
- Promote information exchange and knowledge sharing

Hlava (2011): Developing an eclectic terminology registry
Collaborations

- Editorial Offices
- BARTOC
- coliconc
- Glossarissimo
- PoolParty
- ISKO
- Linked Open Vocabularies (LOV)
# Inter-Active Terminology for Europe

**Interoperability**

## Interoperability of KOS metadata schemas (2016-10-16)

<table>
<thead>
<tr>
<th>View</th>
<th>Voting results</th>
</tr>
</thead>
</table>

### Abstract:

"In IATE you can find EU-specific terminology and jargon, as well as terms from all sorts of areas, such as law, agriculture, Information technology and many others. IATE has 8.4 million terms, including approximately 540 000 abbreviations and 130 000 phrases, and covers all 25 official EU languages. New terms are added every day and the contents are constantly updated. IATE plays a major role in ensuring the quality of the written communication of the EU institutions and bodies."

### Alternative Title: IATE

| Size: | 8.4 million terms, including approximately 540 000 abbreviations and 130 000 phrases (2016-06) |

| Year of Creation: | 1999 |

| Author: | Translation Centre for the Bodies of the European Union |

| VIAF: | Translation Centre for the Bodies of the European Union |

| Link: | [http://iate.europa.eu](http://iate.europa.eu) |

### Topic: European Union

| DDC: | 001 Knowledge |

| ILC: | civil society, communities, groups, people, governments, political power, political units, sovereign states, economies, economic systems, technologies, technological systems, production systems |

| s: | UE: European Union, EU |

| Access: | Free |

| Format: | Online |

| Type: | thesaurus |

### Wikidata:

- [http://www.wikidata.org/entity/Q1520860](http://www.wikidata.org/entity/Q1520860)

### Wikipedia:

- be, da, de, en, es, et, fr, gd, gl, it, lv, lt, pt, ru, uk

### Language:

- Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latin, Lithuanian, Maetisian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish

| Location: | Luxembourg, Unknown, Luxembourg |

| Geo: | POINT (6.1424653 49.6184686) |

| License: | Unknown |

### EuroVoc URIs:


### ISO 639-2 URIs:

Interoperability of KOS metadata schemas (2016-10-16)
NKOS DC Application Profile

Zeng and Žumer (2015): Networked Knowledge Organization Systems Dublin Core Application Profile (NKOS AP) v0.2

- BARTOC is aligned with NKOS AP
- Some NKOS AP fields are not included or applicable [yet?]
- Additional constraints make sense

$\rightarrow$ BARTOC RDF $\subset (\text{NKOS AP } \cup \text{Schema.org } \cup \text{FOAF})$
<http://bartoc.org/en/taxonomy/term/1> a skos:Concept;
   rdfs:label "thesaurus"@en ; skos:prefLabel "thesaurus"@en .

   dc:type <http://bartoc.org/en/taxonomy/term/1>;
   schema:name "IATE", "Inter-Active Terminology for Europe"@en;
   dc:title "IATE";
   dc:creator "Translation Centre for the Bodies of the European Union";
   foaf:maker <http://viaf.org/viaf/167896131>,
   "Translation Centre for the Bodies of the European Union";
   dc:created "1999"^^xsd:gYear;
   nkos:sizeNote
   "8.4 million terms, including approximately 540,000 abbreviations and 130,000 phrases (2016-06)" ;
   dc:description "In IATE you can find EU-specific terminology and jargon, as well as terms from all sorts of areas..."
   dc:identifier <http://iate.europa.eu>;
   dc:relation <http://www.wikidata.org/entity/Q1520860>;
   schema:about
   <http://bartoc.org/en/taxonomy/term/10656>,
   <http://bartoc.org/en/taxonomy/term/11597>,
   <http://bartoc.org/en/taxonomy/term/50220>,
   <http://bartoc.org/en/taxonomy/term/50430>,
   <http://bartoc.org/en/taxonomy/term/49253>,
   dc:language
   <http://bartoc.org/en/taxonomy/term/10746>,
   <http://bartoc.org/en/taxonomy/term/10747>,
   <http://bartoc.org/en/taxonomy/term/10748>,
   <http://bartoc.org/en/taxonomy/term/10749>,
   ...

Interoperability of KOS metadata schemas (2016-10-16)
BARTOC KOS Metadata in RDF

<http://bartoc.org/en/taxonomy/term/1> a skos:Concept; # == <http://w3id.org/nkos/nkostype#thesaurus>
  rdfs:label "thesaurus"@en ; skos:prefLabel "thesaurus"@en.

  dc:type <http://bartoc.org/en/taxonomy/term/1>; # == <http://w3id.org/nkos/nkostype#thesaurus>
  schema:name "IATE", "Inter-Active Terminology for Europe"@en;
  dc:title "IATE";
  dc:creator "Translation Centre for the Bodies of the European Union";
  foaf:maker <http://viaf.org/viaf/167896131>, # => VIAF
    "Translation Centre for the Bodies of the European Union";
  dc:created "1999"^^xsd:gYear;
  nkos:sizeNote
    "8.4 million terms, including approximately 540.000 abbreviations and 130 000 phrases (2016-06)";
  dc:description "In IATE you can find EU-specific terminology and jargon, as well as terms from all sorts ..."
  dc:identifier <http://iate.europa.eu>;
  dc:relation <http://www.wikidata.org/entity/Q1520860>; # => Wikidata
  schema:about
    <http://bartoc.org/en/taxonomy/term/50220>, # == s: civil society, communities, groups, people
    <http://bartoc.org/en/taxonomy/term/50430>, # == t: governments, political power, political units, sovereignty
    <http://bartoc.org/en/taxonomy/term/50770>; # == tUE: European Union, EU
    <http://bartoc.org/en/taxonomy/term/51043>; # == u: economies, economic systems
  dc:language
    <http://bartoc.org/en/taxonomy/term/10746>, # == en (English)
    <http://bartoc.org/en/taxonomy/term/10747>, # == fr (French)
    <http://bartoc.org/en/taxonomy/term/10748>, # == de (German)
    <http://bartoc.org/en/taxonomy/term/10749>, # == es (Spanish)
...
Metadata Schema Conversion

manual input → BARTOC database fields (Drupal)
⇒ Drupal → BARTOC RDFa (NKOS AP & schema.org)
⇒ Wrapper → JSKOS (SKOS & extensions)

Conversions don’t preserve all information (⇔) but normalize

See Ledl and Voß (2016) for conversion details
JSKOS

- based on RDF (SKOS and DC Terms)
  - additional constraints for better normalization
  - support of mappings, concordances, and registries in addition to concept schemes and concepts
- based on JSON-LD
  - simplified and extended with closed world statements

Primary use case: **simple use in web applications**

Specified at [https://gbv.github.io/jskos/](https://gbv.github.io/jskos/)
BARTOC KOS Metadata in JSKOS

{
  "uri": "http://bartoc.org/en/node/20",
  "type": [
    "http://www.w3.org/2004/02/skos/core#ConceptScheme",
    "http://w3id.org/nkos/nkostype#thesaurus"
  ],
  "prefLabel": {
    "en": "Inter-Active Terminology for Europe"
  },
  "notation": [ "IATE" ],
  "created": "1999",
  "creator": [ {
    "uri": "http://viaf.org/viaf/167896131",
    "prefLabel": { "en": "Translation Centre for the Bodies of the European Union" } }
  ],
  "extent": "8.4 million terms, including approximately 540 000 abbreviations and 130 000 phrases (2016-06)",
  "identifier": [ "http://www.wikidata.org/entity/Q1520860" ],
  "url": "http://iate.europa.eu",
  "languages": [ "en", "fr", "de", "es", null ] # null: explicit existence
  "subject": [ {
    "uri": "http://eurovoc.europa.eu/4060",
    "prefLabel": { "en": "European Union" },
  },{
    "uri": "http://dewey.info/class/001/e23/"
    "notation": [ "001" ],
    "prefLabel": { "en": "Knowledge" }
  },
  null # null: explicit existence
}
JSKOS-API Wrapper

- Implementation and demo: https://github.com/gbv/jskos-php-examples
- Straightfoward mapping from RDF to SKOS
- API specified at https://gbv.github.io/elma/
# Publications

**BARTOC database dumps and reports**

This page contains database dumps of [BARTOC.org](http://www.bartoc.org) converted to [JSKOS format](http://www.jskos.org). All data is made available under [CC Zero](http://creativecommons.org/publicdomain/zero). 

- concept schemes in JSKOS (1905, last updated 2016-10-13 02:00)
  - download all (NDJSON)
- terminology registries in JSKOS (75, last updated 2016-10-13 02:23)
  - download all (NDJSON)
- data quality reports of concept schemes in JSKOS

Please report errors in BARTOC to the [BARTOC editors](http://www.bartoc.org) and errors in the database dumps at [this issue tracker](http://www.bartoc.org).

## Concept schemes report

<table>
<thead>
<tr>
<th>id</th>
<th>notation</th>
<th>name</th>
<th>created</th>
<th>license</th>
<th>Wikidata</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>AGIFT</td>
<td>Australian Governments' Interactive Functions Thesaurus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>AHSP</td>
<td>Descrição Arquivística pt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>AIM25</td>
<td>Thesaurus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>APAIS</td>
<td>Thesaurus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>CSH</td>
<td>Canadian Subject Headings</td>
<td></td>
<td></td>
<td>Q5030602</td>
</tr>
<tr>
<td>12</td>
<td>ALearn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>CSPS</td>
<td>Controlled Vocabulary</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

- > 1,900 KOSs (and 70 KOS registries) in BARTOC
- RDF Metadata Schema: NKOS AP & schema.org & FOAF
- Controlled vocabularies: EuroVoc, DDC, Wikidata
  ISO 639-2 languages, NKOS KOS Types, ILC...
- JSKOS format: SKOS & JSON-LD & constraints/extensions
- BARTOC dumps available as Open Data
- Contributions are welcome!
See also http://bartoc.org/ and https://coli-conc.gbv.de/publications.

