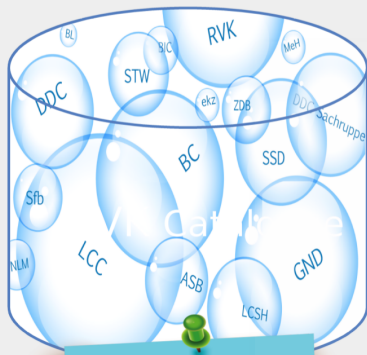


# Describing Knowledge Organization Systems in BARTOC and JSKOS

Andreas Ledl, Jakob Voß, Uma Balakrishnan

TKE 2016, Copenhagen, 2016-06-23

# Project coli-conc

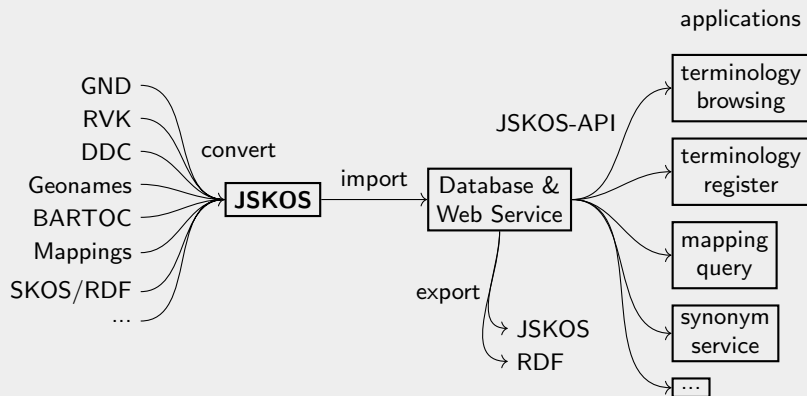


## Aim

Semi-automatic creation of mappings between library knowledge organisation systems

Classification systems in German speaking regions	
Universal Classification Systems	No. of classes
UDC (Universal Decimal Classification)	ca. 65.000 classes (English version)
DDC (Dewey Decimal Classification)	over 44.000 classes with 10 main classes
RVK (Regensburg Classification)	850.000 classes with 33 main classes
BC (Basic Classification)	2100 classes with 89 main classes
LCC (Library of Congress Classification)	21 main classes
Subject classification	No. of classes
DDC-Sachgruppen der DNB	10 main classes with 94 subclasses
MSC (Mathematics Subject Classification)	87 main classes
PACS (Physics and Astronomy Classification Scheme)	10 main classes
FKDigBib (Subject classification for digital library)	10 main classes
KfM (Classification for music library)	ca. 800 classes
Subject Classification at the Universities	No. of classes
TUM-classification (Science and technology classification of the TU Munich)	52 classes each with 999 notations
Subject classification of the University library Duesseldorf	45 classes
Bremer classification of the State and University library Bremen	ca. 57 main classes
GOK (Goettingen Online Classification)	ca. 33 main classes
Standard-Thesaurus Wirtschaft von der ZWB	6.000 Terms and notations
Subject classification University library Trier	36 main classes
Technical University Dortmund	28 main classes
University library Paderborn	26 main classes
University library Marburg	35 main classes
University library Bonn	24 main classes
University library Heidelberg	22 main classes
Subject classification and nomenclature of individual languages	23 main classes
Library of the Institute of General Linguistics at the Uni Münster	
Subject Classification at the public libraries	No. of classes
SEB (Scheme for protestant libraries)	
SKB-E (Scheme for catholic public libraries)	
KfKJ (Scheme for children and youth libraries)	Less than 1.000 classes
ASB (General classification for public libraries)	ca. 2.200 classes with 23 main classes
ÖSÖB (Austrian classification for public libraries)	
SfB (Classification for libraries)	ca 14.400 classes with 30 main classes
KAB (Classification for general libraries)	ca. 2.700 classes
SSD (Classification of the city library Duisburg)	
ESSB (Single classification for South Tyrolean)	16 main classes

# System overview



# JSKOS data format

Requirement: uniform access to KOS data

- ▶ based on RDF (SKOS & Dublin Core)  
... but more restrictive (unification)
- ▶ based on JSON-LD  
... extended by language ranges and closed world statements

main goal: **easy use of KOS data in web applications**

Specified at <https://gbv.github.io/jskos/>

# JSKOS beyond JSON-LD: closed world statements

narrower concepts may exist (open world assumption)

▶ { }

∃: narrower concepts exist (but not included here)

▶ { "narrower": [ null ] }

⧸: no narrower concepts exist

▶ { "narrower": [] }

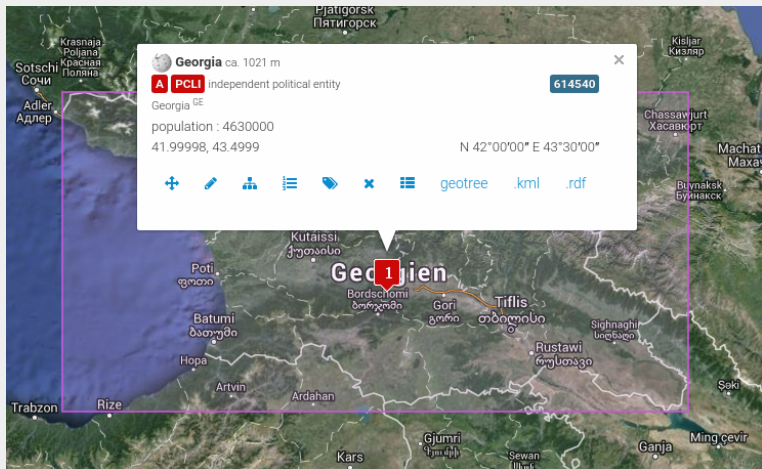
language ranges state that more labels in other languages exist

▶ { "en": "english label", "-": "..." }

# JSKOS beyond SKOS: data and object types

data types	object types
URI	Concept
URL	Concept Type
date	Concept Scheme
list	Registry
set	Concordance
language range	Concept Mapping
language map	Concept Bundle

# Concept example from GeoNames



<http://sws.geonames.org/614540/>

# JSKOS concept example from GeoNames (abbreviated)

```
{
  "type": ["http://www.w3.org/2004/02/skos/core#Concept"],
  "uri": "http://sws.geonames.org/614540/",
  "inScheme": [
    { "uri": "http://bartoc.org/en/node/15" }
  ],
  "prefLabel": {
    "en": "Georgia",
    "fr": "Géorgie",
    "de": "Georgien",
    "-": "..."
  },
  "broader": [
    { "uri": "http://sws.geonames.org/6255147/" }
  ],
  ...
}
```



# BARTOC in JSKOS

KOS can also be expressed in JSKOS

- ▶ Mapping of BARTOC fields to JSKOS
- ▶ BARTOC JSKOS export available:  
<https://github.com/gbv/bartoc-dumps/>

# BARTOC example: Eurovoc

## Multilingual Thesaurus of the European Union

View

Voting results

### Abstract:

"EuroVoc is a multilingual thesaurus which was originally built up specifically for processing the documentary information of the EU institutions.

It is a multi-disciplinary thesaurus covering fields which are sufficiently wide-ranging to encompass both Community and national points of view, with a certain emphasis on parliamentary activities. EuroVoc is a controlled set of vocabulary which can be used outside the EU institutions, particularly by parliaments.

The aim of the thesaurus is to provide the information management and dissemination services with a coherent indexing tool for the effective management of their documentary resources and to enable users to carry out documentary searches using controlled vocabulary.

The EuroVoc ontology is an extension of SKOS (Simple Knowledge Organization System) - W3C recommendation, including appendix B, SKOS eXtension for Labels (SKOS-XL)."

**Alternative Title:** EuroVoc

**Author:** Publications Office of the European Union

**VIAC:** Publications Office of the European Union

**Link:** <http://eurovoc.europa.eu/>

**Topic:** European Union politics

**DDC:** 001 Knowledge

**Access:** Registered Free

**Format:** Online OWL Printed RDF SKOS XML

**Type:** ontology thesaurus

**Wikidata:** <http://www.wikidata.org/entity/Q1370467>

**Wikipedia:** [ca](#) [de](#) [en](#) [eo](#) [hr](#) [lt](#) [nl](#) [pt](#) [ro](#) [uk](#)

**Language:**

Bulgarian Catalan Croatian Czech Danish Dutch English Estonian Finnish French German  
Greek Hungarian Italian Latvian Lithuanian Macedonian Maltese Polish Portuguese Romanian  
Serbian Slovak Slovenian Spanish Swedish

### EuroVoc URIs

<http://eurovoc.europa.eu/4060>

<http://eurovoc.europa.eu/4704>

### ISO 639-2 URIs

<http://id.loc.gov/vocabulary/iso639-2/bul>

<http://id.loc.gov/vocabulary/iso639-2/cat>

<http://id.loc.gov/vocabulary/iso639-2/hrv>

<http://id.loc.gov/vocabulary/iso639-2/cze>

<http://id.loc.gov/vocabulary/iso639-2/dan>

<http://id.loc.gov/vocabulary/iso639-2/dut>

<http://id.loc.gov/vocabulary/iso639-2/eng>

<http://id.loc.gov/vocabulary/iso639-2/est>

<http://id.loc.gov/vocabulary/iso639-2/fin>

<http://id.loc.gov/vocabulary/iso639-2/fre>

<http://id.loc.gov/vocabulary/iso639-2/ger>

<http://id.loc.gov/vocabulary/iso639-2/gre>

# JSKOS BARTOC example: Eurovoc record (abbreviated)

```
{  
  "type": ["http://www.w3.org/2004/02/skos/core#ConceptScheme"],  
  "uri": "http://bartoc.org/en/node/15",  
  "url": "http://eurovoc.europa.eu/",  
  "identifier": [  
    "http://www.wikidata.org/entity/Q1370467"  
  ],  
  "prefLabel": {  
    "en": "Multilingual Thesaurus of the European Union"  
  },  
  "altLabel": { "en": [ "EuroVoc" ] },  
  "languages": [ "bg", "ca", "cs", "da", "de", "el",  
    "en", "es", "et", "fi", "fr",  
    ...  
  ]  
}
```


- ▶ HTTP API to query KOS data in JSKOS
- ▶ Entity Lookup Microservice API (ELMA)  
as most relevant subset of JSKOS-API

`entity lookup` ?uri=<http://sws.geonames.org/614540/>

`entity search` ?search=Georgia (labels, notations, notes...)

- ▶ more use cases will be covered by full JSKOS-API
- ▶ public demo of terminology wrappers:  
<https://jskos-php-examples.herokuapp.com/>  
(BARTOC, Wikidata, GND, ORCID, GeoNames, VIAF ...)

## Normdatendienst

Normdatei Dewey-Dezimalklassifikation (DDC) 

Suche

612.112

- Technik, Medizin, angewandte Wissenschaften 6
- Medizin und Gesundheit 61
- Humanphysiologie 612
- Blut und Kreislauf 612.1
- Blut 612.11
- Leukozyten (Weiße Blutkörperchen) 612.112
- Biophysik 612.1122
- Anzahl und Auszählung 612.1127

# Example: Dewey Decimal Classification (DCC)

DDC in BARTOC (<http://bartoc.org/en/node/241>) → concept scheme

```
{  
  "type": ["http://www.w3.org/2004/02/skos/core#ConceptScheme"],  
  "uri": "http://bartoc.org/en/node/241",  
  "prefLabel": { "en": "Dewey Decimal Classification" }  
}
```

DDC in MARCXML (MARC 21 format for classification data) → concepts

```
{  
  "type": ["http://www.w3.org/2004/02/skos/core#Concept"],  
  "uri": "http://dewey.info/class/612.112",  
  "prefLabel": { "en": "leucocytes" },  
  "notation": [ "612.112" ],  
  "inScheme": [ { "uri": "http://bartoc.org/en/node/241" } ]  
}
```

# Mapping Database

[coli-conc](#)[About](#)[Terminologies](#)[Concordances](#)[Cocoda prototype](#)[Publications](#)[Contact](#)

## Concordances

This page provides an early preview of [coli-conc mapping database](#).

### Search Mappings

Source

Target

Creator

Found 2 mappings.

Source Scheme	Source Concept	Target Scheme	Target Concept(s)	Creator <a href="#">!i</a>
<a href="#">DDC</a>	<a href="#">612.112</a>	<a href="#">RVK</a>	<a href="#">WW 8844</a>	VZG
<a href="#">DDC</a>	<a href="#">612.112</a>	<a href="#">RVK</a>	<a href="#">WW 8840</a>	VZG

## Documentation

Coli-conc mapping database is accessible [JSKOS-API](#) at <https://coli-conc.gbv.de/cocoda/api>. See [GitHub repository](#) for source code and technical documentation.

coli-conc is a project of the head office of GBV – [Verbundzentrale des GBV \(VZG\)](#) – funded by German Research Foundation (DFG)

[@coli\\_conc](#)

# JSKOS mapping example

```
{
  "creator": { "notation": [ "VZG" ] },
  "mappingMethod": "...",
  "fromScheme": {
    "uri": "http://bartoc.org/en/node/241",
    "notation": [ "DDC" ]
  },
  "from": { "members" : [
    { "uri": "http://dewey.info/class/612.112" }
  ] },
  "toScheme" : {
    "uri": "http://bartoc.org/en/node/533",
    "notation": [ "RVK" ]
  },
  "to": { "members" : [
    { "uri": "http://uri.gbv.de/terminology/rvk/WW8844" }
  ] }
}
```



## Other possible applications

- ▶ semantic tagging of documents with identified concepts:  
<http://sws.geonames.org/614540/> ⇒ “Georgia”
- ▶ statistical analysis of terminologies:  
compare size, hierarchy, density of links. . .
- ▶ enrichment of concepts from multiple KOS
- ▶ . . .

*foster the visibility, availability and usefulness of KOS in general*

# Thanks! Questions?

- ▶ Follow us on Twitter
  - ▶ @BARTOC\_UBBasel
  - ▶ @coli\_conc
- ▶ More information
  - ▶ <http://bartoc.org/>
  - ▶ <https://coli-conc.gbv.de/>
  - ▶ <https://gbv.github.io/jskos/>
  - ▶ <https://gbv.github.io/elma/>

*Feedback and contribution is appreciated!*