

10 May 2019
EDUG 2019, Stockholm

Cocoda Mapping Tool

Stefan Peters
Verbundzentrale des GBV (VZG)
stefan.peters@gbv.de

What is Cocoda?

- Cocoda is a web-based tool for semi-automatic creation of mappings between different knowledge organization systems (KOS).
- Goal: Move away from Excel spreadsheets and multiple browser tabs for research, and integrate everything into one interface.
- Primarily developed with library classifications (DDC, RVK, BK, GND, ...) in mind, but supports all kinds of other KOS as well.

Background

- Developed by the GBV Head Office (VZG) in Göttingen, Germany.
- Part of Project coli-conc:
<https://coli-conc.gbv.de>
- The mapping tool and its background services are available as open source repositories on GitHub.
- The data is openly available through APIs.

Data Sources

- Data sources use the JSKOS format (SKOS in JSON-LD, <https://gbv.github.io/jskos/>).
- In Cocoda, different sources are configurable for
 - vocabularies
 - concepts
 - existing mappings
 - mapping recommendations

Mappings

- Project coli-conc collected existing concordances with around 385,000 mappings.
- Additionally, mapping data from Wikidata is available.
- First use in production by Project KENOM (<https://www.kenom.de>):
 - Mapping of local systems to GND, Wikidata, Nomisma, and others.
 - ~28,000 mappings created since November 2018.

Multilanguage Support

- Cocoda has support for multiple interface languages, currently English and German.
- New languages can be added easily (contact us if you'd like to add Swedish 😊).
- Supports vocabulary data in multiple languages, e.g. Wikidata.
 - Currently, most of the available vocabularies in Cocoda are in German.

Data Conversion

- Add your own data etc. (TODO)

Live Demo

- <https://coli-conc.gbv.de/cocoda/app/>
- Note: DDC in English is currently not available in the release version of the tool, it is only used here for demo purposes.



Future Plans

- Add more KOS.
- Offer tools to easily convert one's own systems into JSKOS format.
- Improve the tool (usability, tutorial, new features).
 - Already collected many ideas:
<https://github.com/gbv/cocoda/issues>
 - e.g.: coli-ana integration (see next slide)

Example:

coli-ana in Cocoda

- Decomposition of synthesized DDC notations right in Cocoda:

 700.90440747471 

Info	Search Links	coli-ana
700.90440747471		
7		Arts & recreation
70		Arts
700		The arts
700.904		Modern arts
T1--0904		20th century, 1900-1999
T1--09044		1940-1949
T1--074		Museums, collections, exhibits
T2--7		North America
T2--74		Northeastern United States (New England and Middle Atlantic states)
T2--747-749		Middle Atlantic states
T2--747		New York (State)
T2--7471		New York Metropolitan Area

More Information

- Slides: <https://exo.pm/edug2019/slides.pdf>
- Project homepage: <https://coli-conc.gbv.de>
- GitHub repository for Cocoda:
<https://github.com/gbv/cocoda>
- Cocoda documentation and manual:
<https://gbv.github.io/cocoda/>