

Webseminar coli-conc

Einführung in die Dienste und das Mapping-Tool Cocoda

22. Januar 2021

Uma Balakrishnan, Jakob Voß



Was ist coli-conc?

coli-conc.gbv.de

- **Dienstleistung der Verbundzentrale des GBV (VZG)**
 - coli-conc bietet eine Infrastruktur zur Bereitstellung und Verwaltung von Mappings und Konkordanzen zwischen Wissensorganisationssystemen (KOS).
- **Entwicklung - Im Rahmen eines DFG-Projektes von der VZG**
 - Förderung 2015-2023, laufende Fortführung

Was ist KOS-Mapping

Mapping als Verb

KOS Mapping ist ein Prozess, der eine Relation zwischen zwei Begriffen herstellt.

Konzepte einer Begriffswelt/eines Systems werden mit dem Konzept eines anderen System in Beziehung gesetzt.

Was ist KOS Mapping

Mapping als Substantiv

KOS Mapping ist ein Tripel (S,T,R) mit Begriffen s und t:

$s \in \{\text{Quellsystem } S\}$ (source system)

$t \in \{\text{Zielsystem } T\}$ (target system)

$r: S \rightarrow T \in \{\text{Relation } R\}$ (r : Beziehung zwischen Begriffen zweier Systeme, z. B. zwischen zwei bedeutungsgleichen Begriffen [exact match], Oberbegriffsbeziehung [s is broader term of t], Unterbegriffsbeziehung [s is narrower term of t])

“In der Mathematik ist eine Funktion (lat. *functio*) oder Abbildung eine Beziehung (Relation) zwischen zwei Mengen, die jedem Element der einen Menge (...) genau ein Element der anderen Menge (...) zuordnet.”

[https://de.wikipedia.org/wiki/Funktion_\(Mathematik\)](https://de.wikipedia.org/wiki/Funktion_(Mathematik))

Was ist eine Konkordanz?

Die Menge aller Mappings von
einem Quell- zu einem Zielsystem

Warum brauchen wir Mappings

- Verbesserung der Interoperabilität zwischen KOS
 - Mappings verbinden und bilden ein Netzwerk der KOS
- Unterstützung bei der Sacherschließung
 - Mappings übersetzen die Begriffe in das Vokabular des Zielsystem (manuell, automatisch)

Warum brauchen wir Mappings

- Unterstützung bei der Suche
 - Ermöglichen von Suchexpansion (intellektuell oder automatisch) für Freitextsuche
 - Ermöglichen einheitlicher Suche durch Überwindung der heterogen erschlossenen Datenbestände
- Erhöhung der Qualität des Retrievals

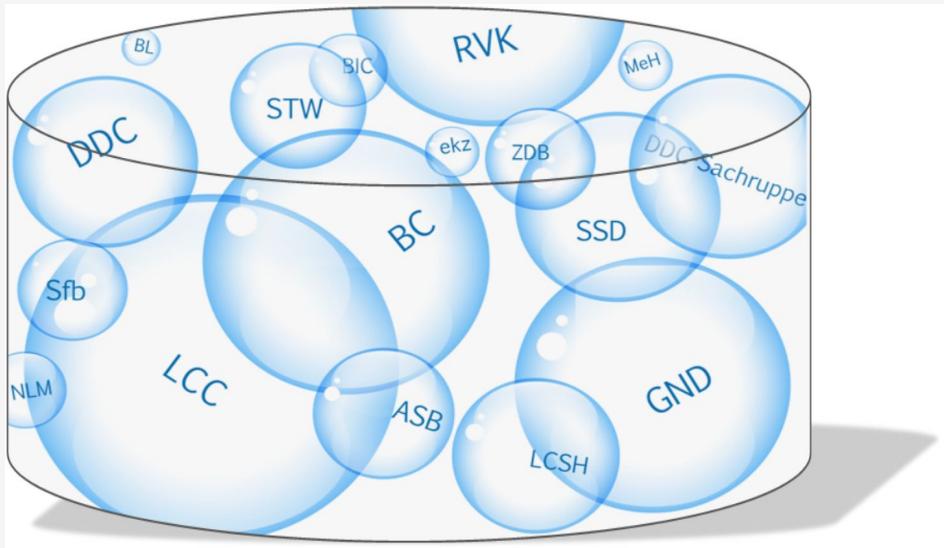
KOS Mapping sind schwierig

- Die KOS sind:
- Sprachspezifisch
- Kulturspezifisch
- Domänenspezifisch
- Unterschiede in der Struktur
- Umfangreich und
- Fachkenntnis von Experten notwendig

Mapping Methode

- One-to-one Mapping (uni- o. bidirektional)
- One-to-Many Mapping (uni- o. bidirektional)
- Many-to-Many Mapping (uni- o. bidirektional)
- Hub Modell (für zwei o. mehr KOS)

coli-conc Ausgangslage



- Hohe Diversität an KOS
- Geringere Verfügbarkeit an konkordanzen
- Wunsch nach Konkordanz zwischen RVK & DDC groß
- Bedarf an einer zentralen Plattform für einfacheren Zugriff und Nachnutzung auf KOS und Konkordanzen
- Automatische Verfahren unzureichend
- Kein Goldstandard für Speicherung, Bereitstellung und Pflege der Konkordanzen

Schwerpunkt des coli-conc Projektes

1. Anwendung auf KOS im bibliothekarischen Bereich
2. Computer-gestützt (Semi-automatisch)
3. Verfahren und Dienste für einfacheren Austausch, Pflege
und Nutzung der KOS und Konkordanzen

Ziele

- Interoperabilität zwischen KOS
- Überwindung von Heterogenität in den Datenbanken
- Effektivere Erstellung und Pflege von Konkordanzen zwischen bibliothekarischen KOS
- Verbesserung der Qualität vorhandener Konkordanzen
- Förderung des Austauschs und der Nutzung von Konkordanzen und KOS
- Unterstützung der Fachreferenten bei der Sacherschließung
- den K10plus Katalog mit Mappings automatisch anreichern

Partner



coli-conc Dienstleistung

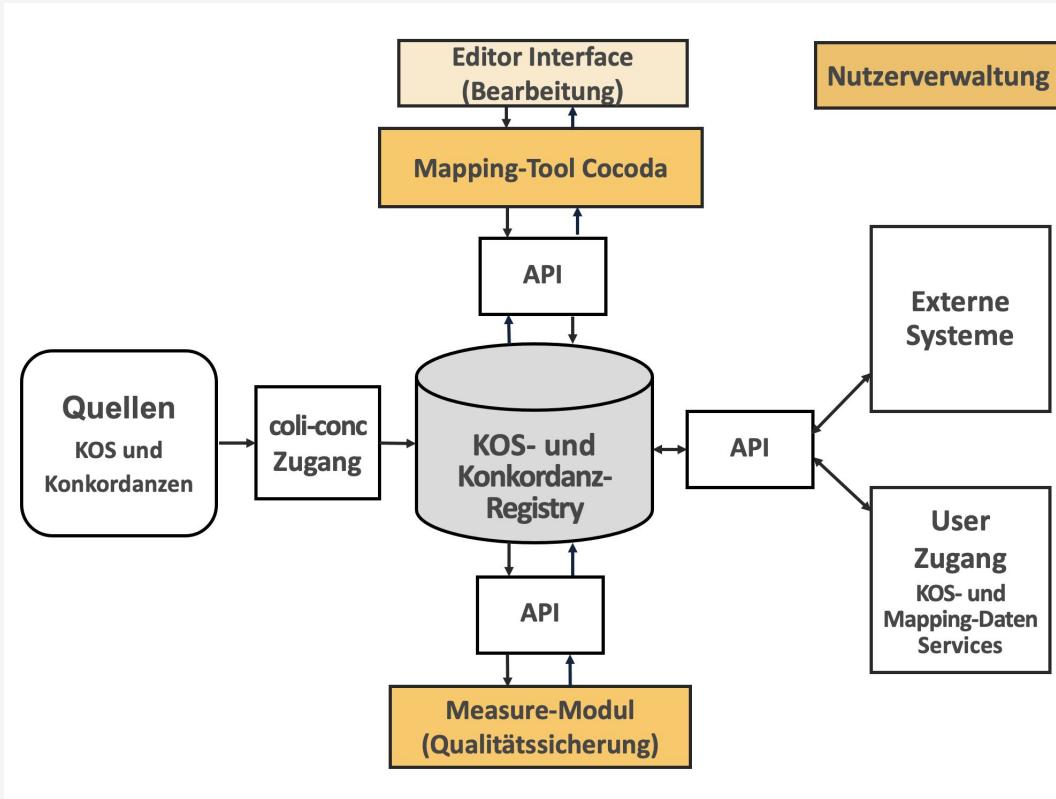
- Einheitlicher Zugriff auf Wissensorganisationssysteme und ihre Metadaten in dem einheitlichen JSKOS-Format
- Eine offene Datenbank zur Speicherung und zum Austausch von KOS und Mappings
- Ein webbasiertes Tool zur vereinfachten Erstellung, Verwaltung und Bewertung von Mappings
- Freie Software zum Import und Export von Daten zu Mappings und Wissensorganisationssystemen
- Anbindung an Portale und Software für Sacherschließung

Hauptbestandteile von coli-conc

- **JSKOS-Datenformat:** <https://gbv.github.io/jskos/jskos.html>
- **JSKOS-API für KOS, Konzepte, Konkordanzen und Mappings**
Implementiert durch die Open Source Software JSKOS-Server
- **KOS-Registry:** Metadaten von bibliothekarischen KOS
wird durch BARTOC ersetzt, welches kürzlich von der VZG übernommen wurde
- **Konkordanz-Registry:** Datenbank mit ca. 370.000 Mappings aus verschiedenen Quellen und zwischen verschiedenen KOS
- **KOS-Datenquellen:** Zugriff auf KOS-Konzepte via Browsing oder Suche, entweder in unserer Datenbank oder per externer Schnittstelle
- **Mapping-Tool Cocoda:** Das Herzstück von coli-conc für die semi-automatische Erstellung von Mappings zwischen bibliothekarischen KOS

Übersicht über die Infrastruktur

- modulare Systemarchitektur
- verschiedene Komponenten können unabhängig voneinander genutzt und vorhandene Programme eingebunden werden
- Grundlage ist das gemeinsame Datenformat JSKOS und JSKOS-API



Cocoda Mapping Tool

Imprint Privacy Policy Feedback Manual ⚒ ★ Stefan Peters

Mapping Editor

RVK DDC

CA - CK Philosophie → 100 Philosophie & Psychologie

Concordance Registry: not saved

Stefan Peters

Mappings

Concordance Registry

DDC	RVK	CA - CK Philosophie	VZG	±0
100 Philosophie & Psychologie	RVK	CA - CK Philosophie	VZG	±0
100 Philosophie, Parapsychologie und Okultismus, Psychologie	RVK	CA - CK Philosophie	VZG	2013
108 Philosophie	RVK	CA - CK Philosophie	Gerald Ja...	2020-05-27

Local Wikidata-Mappings

Recommendations

CR coli-conc Recommendations

RVK	CA - CK Philosophie	DDC	113.8 Philosophie des Lebens
RVK	CA - CK Philosophie	DDC	801 Philosophie und Theorie
RVK	CA - CK Philosophie	DDC	701 Philosophie und Theorie der bildenden und angewandten Kunst
RVK	CA - CK Philosophie	DDC	601 Philosophie und Theorie
RVK	CA - CK Philosophie	DDC	501 Philosophie und Theorie
RVK	CA - CK Philosophie	DDC	401 Philosophie und Theorie; internationale Sprachen
RVK	CA - CK Philosophie	DDC	335.401 Philosophie
RVK	CA - CK Philosophie	DDC	10 Philosophie

ccmapper Mapping Recommendations

RVK	CA - CK Philosophie	DDC	3C-384 Philosophische Themen
-----	---------------------	-----	------------------------------

Tree View

A Allgemeines
B Theologie und Religionswissenschaften
CA - CK Philosophie
CL - CZ Psychologie
D Pädagogik
E Allgemeine und vergleichende Sprach- und Literaturwissenschaft. Indogermanistik, Ältereueuropäische Sprachen und Literaturen
F Klassische Philologie. Byzantinistik, Mittellateinische und Neugriechische Philologie. Niederdeutsch, Skandinavistik
G Germanistik. Niederlandistik, Amerikanistik
H Anglistik. Amerikanistik
I Romanistik
K Slawistik
LA - LC Ethnologie
LD - LG Klassische Archäologie
LD,LH - LO Kunstgeschichte
LP - LY Musikwissenschaft
MA - ML Politologie
MN - MS Soziologie
MT Gesundheitswissenschaften

Tree View

000 Informatik, Informationswissenschaft & allgemeine Werke
100 Philosophie & Psychologie
200 Religion
300 Sozialwissenschaften
400 Sprache
500 Naturwissenschaften
600 Technik, Medizin, angewandte Wissenschaften
700 Künste und Unterhaltung
800 Literatur
900 Geschichte und Geografie
T1-0 Hilfstafel 1. Standardschlüssel
T2-0 Hilfstafel 2. Geographische Gebiete, Zeitabschnitte, Biografien
T3-0 Hilfstafel 3. Schlüssel für Künste, für einzelne Literaturen, für einzelne literarische Gattungen
T3A-0 Hilfstafel 3A. Schlüssel für Werke von oder über Einzelauteuren
T3B-0 Hilfstafel 3B. Schlüssel für Werke von mehr als einem oder über mehr als einen Autor
T3C-0 Hilfstafel 3C. Notationen zum Anhängen aufzunehmen von Anweisungen

Mapping-Tool Cocoda

öffentlicht zugänglich unter:

<https://coli-conc.gbv.de/cocoda/app/>

- benötigt modernen Browser
- Englisch und Deutsch
- Zugriff auf ~170 KOS,
- 27 Konkordanzen und ~370.000 Mappings
- Zum Speichern von Mappings in unsere Datenbank wird eine ORCID oder ein Account bei Wikidata benötigt. Speicherung lokal im Browser ist möglich.

Cocoda live Demo

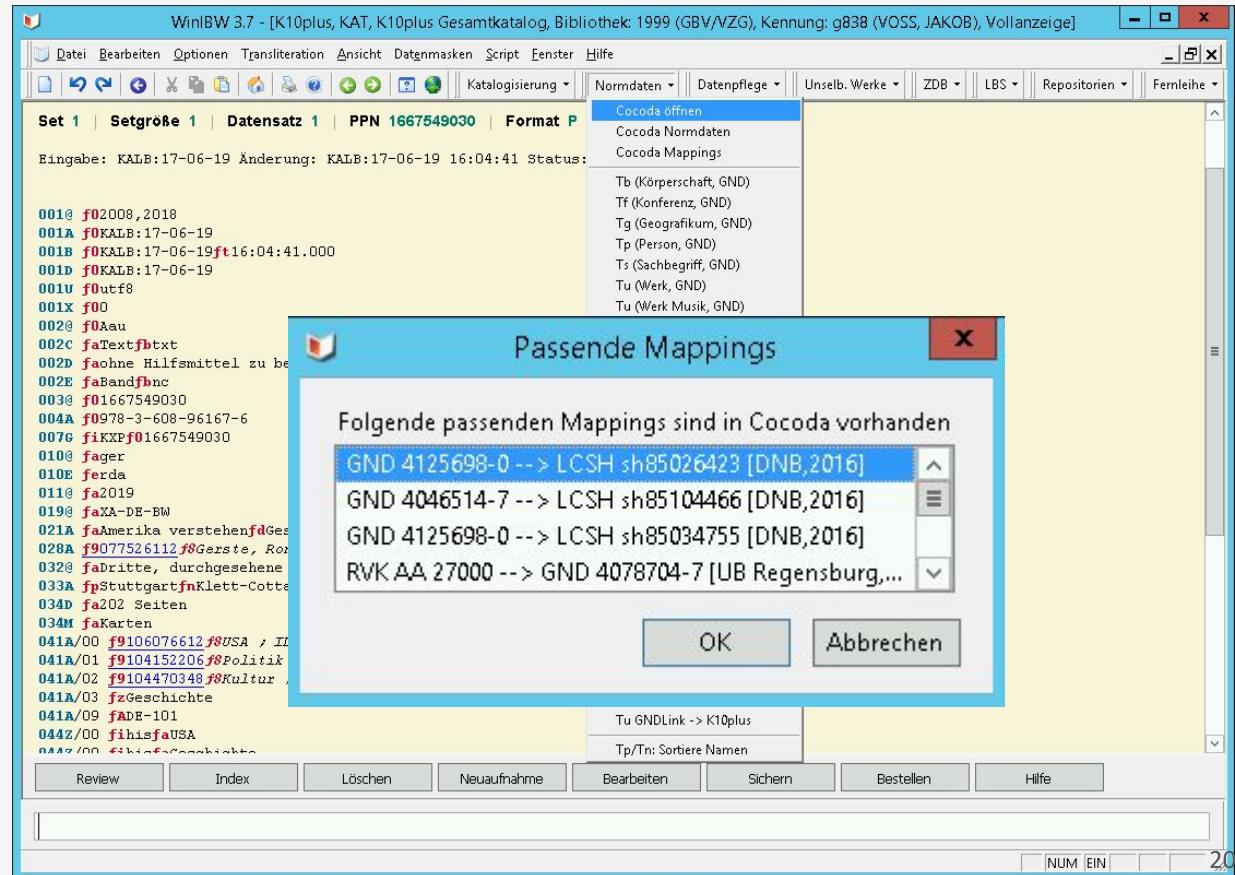
Anbindung an Portale und Anreicherung

Anbindung an WinIBW und DA-3

WinIBW: Proof of Concept zum
Ausprobieren

<https://github.com/gbv/cocoda-winibw>

DA-3: Anbindung geplant



coli-rich: Automatische Anreicherung in K10plus

- Vorhandene Sacherschließung → **Mappings** → Weitere Sacherschließung
- Beispiel

045F=5010 \$a549 DDC 549 Mineralogie

DDC 549 → BK 38.30 vorhandenes Mapping

045Q/01=5301 zusätzliches PICA-Feld mit Quellenangabe
\$a38.30

\$Acoli-conc DDC-RVK

\$A<https://coli-conc.gbv.de/api/mappings/af8ac88b-f7ab-427a-8e06-9e091d281bcd>

- Ausprobieren unter <https://coli-conc.gbv.de/coli-rich/>

Vielen Dank!

Webseite: <https://coli-conc.gbv.de>

Uma Balakrishnan (Projektleitung): balakrishnan@gbv.de

Jakob Voß (Technische Koordination): voss@gbv.de

Stefan Peters (Software-Entwicklung): peters@gbv.de

Twitter: [@coli_conc](https://twitter.com/coli_conc)

Interesse an Hilfskraftstelle
und/oder Abschlussarbeit?

Bitte melden!