

coli-conc

Einführung in die Dienste und das Mapping- Tool Cocoda

25. Januar 2022

Uma Balakrishnan, Stefan Peters





CC-BY-SA Jônatas Cunha <https://w.wiki/4S8h>

Was ist coli-conc?

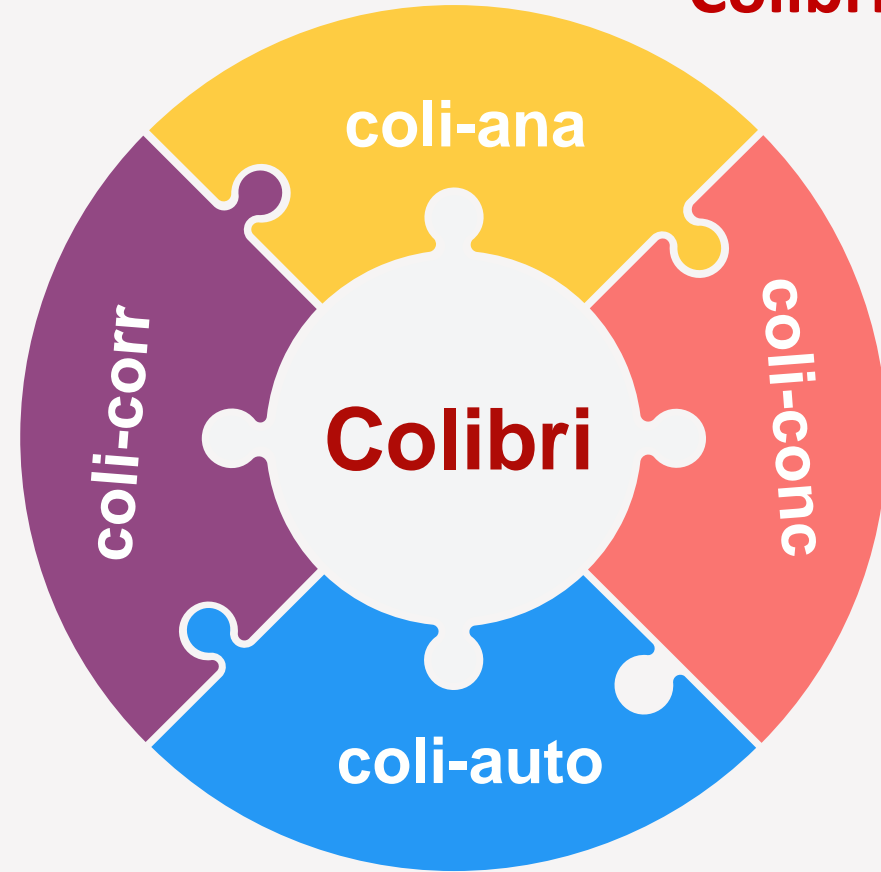
coli-conc.gbv.de

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

Colibri Projekte



CC-BY-SA Jónatas Cunha <https://w.wiki/4S8h>



- coli-ana: automatic analysis
- coli-conc: concordances
- coli-corr: correctness
- coli-auto: automatic classification

Was ist eine 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 Systems in Beziehung gesetzt.



Was ist eine 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])

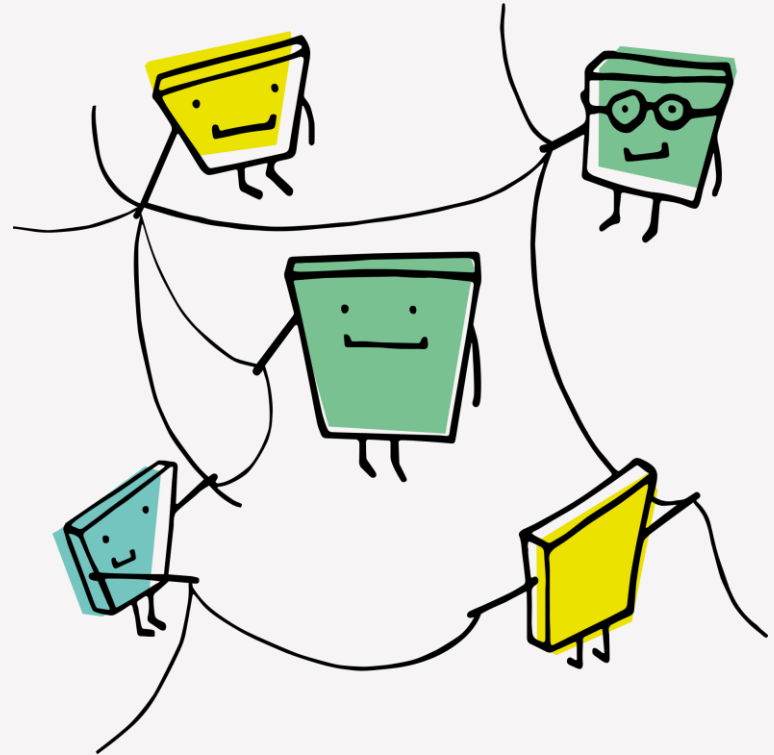
“In der Mathematik ist eine Funktion (lat. functio) oder Abbildung eine Beziehung (Relation) zwischen zwei Mengen, die jedem Element der einen Menge (...) genau ein Element der anderen Menge (...) zuordnet.”

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



Was ist eine Konkordanz?

Eine Menge aller Mappings von einem Quellsystem 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)

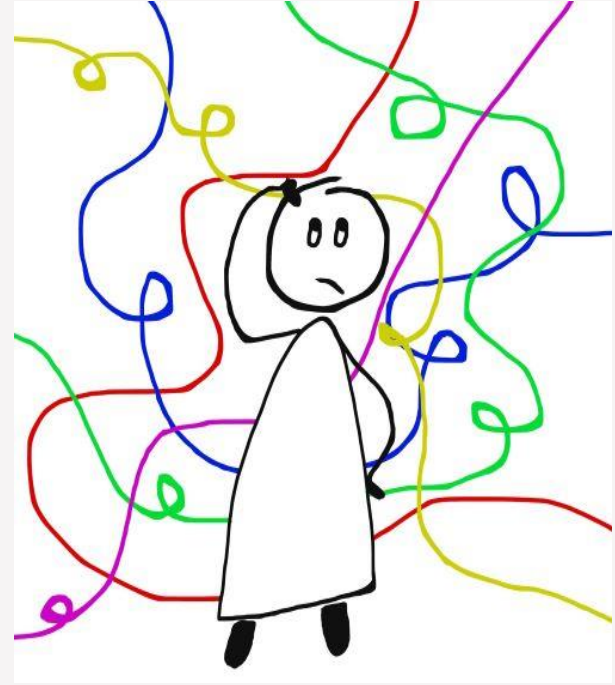
Warum brauchen wir Mappings?

- 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

Warum Mapping ist schwierig?

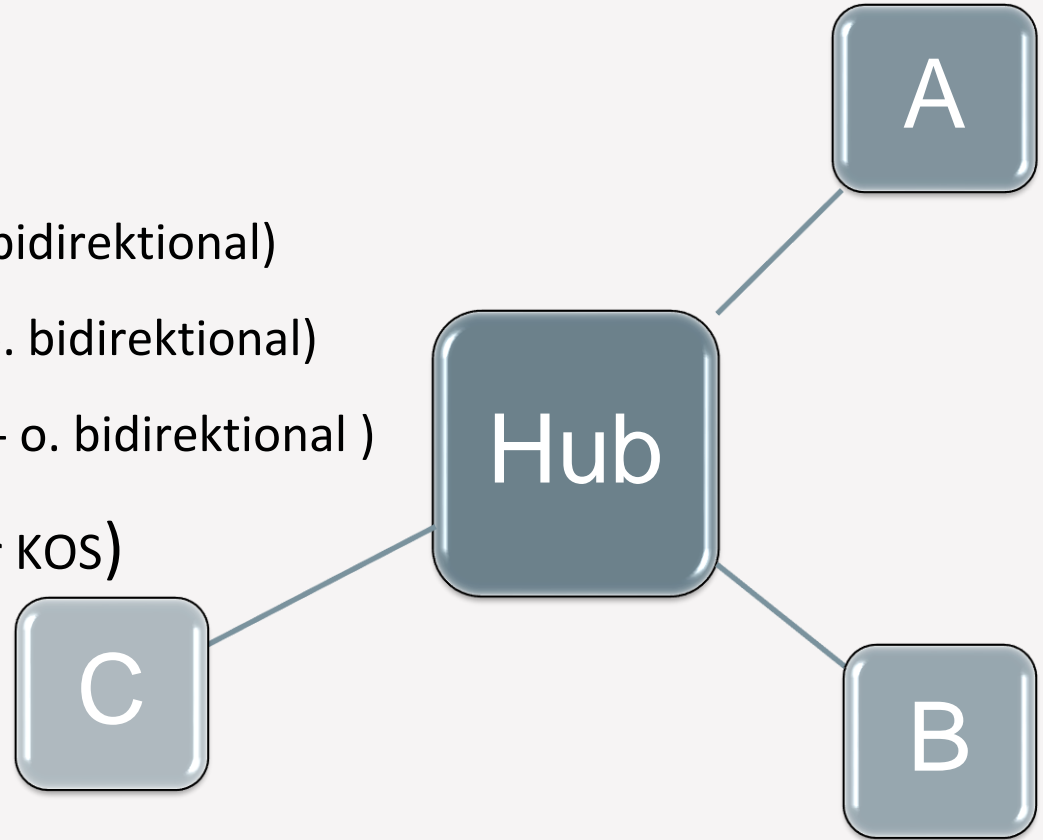
Die KOS sind:

- sprachspezifisch
- kulturspezifisch
- domänenspezifisch
- unterschiedlich in der Struktur
- umfangreich
- benötigen Fachkenntnisse von Expert*innen



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. Computergestützt (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 kollaborative 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

Projektpartner



coli-conc Dienstleistung

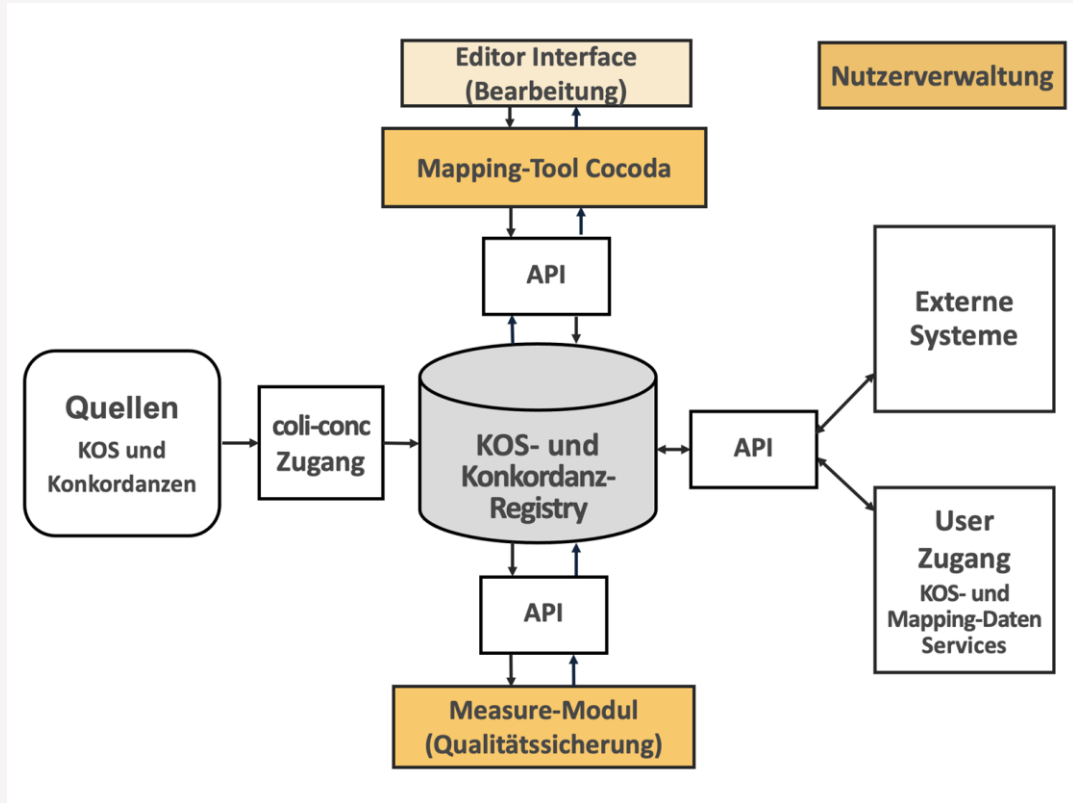
- **Einheitlicher Zugriff** auf Wissensorganisationssysteme und ihre Metadaten in dem einheitlichen JSKOS-Format
- Eine **offene Datenbank** zur Speicherung und zum Austausch von KOS und Mappings
- Ein **webbasiertes Tool** zur vereinfachten Erstellung, Verwaltung und Bewertung von Mappings
- **Freie Software** zum Import und Export von Daten zu Mappings und Wissensorganisationssystemen
- **Anbindung** an Portale und Software für die Unterstützung von 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-Register** : Metadaten von bibliothekarischen KOS
wird durch den **BARTOC – Dienst** ersetzt, welches kürzlich von der VZG übernommen wurde
- **Konkordanz-Register:** Datenbank mit ca. 400.000 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 in vorhandene Programme eingebunden werden
- Grundlage ist das gemeinsame Datenformat JSKOS und die JSKOS-API



Mapping-Tool Cocoda

öffentlich zugänglich unter:

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

- benötigt modernen Browser
- Englisch und Deutsch
- Zugriff auf ~180 KOS, ca. 30 Konkordanzen und ~400.000 Mappings
- Zum Speichern von Mappings in unsere Datenbank wird eine ORCID oder ein Account bei Wikidata benötigt. Speicherung lokal im Browser ist möglich.

The screenshot displays the Cocoda Mapping Tool interface. At the top, the 'Mapping Editor' is active, showing a mapping from 'CA - CK Philosophie' to '100 Philosophie & Psychologie'. Below this, a table lists 'Concordances' with columns for source, target, and date. The 'Tree View' on the left shows a hierarchical structure of knowledge objects (KOS), with 'CA - CK Philosophie' selected. The right sidebar shows the 'DDC Dewey Decimal Classification' and a list of available mappings.

Mapping Editor

RVK → DDC

CA - CK Philosophie → 100 Philosophie & Psychologie

Concordance Registry: not saved

Stefan Peters

Concordances

Source	Target	Date	Score
DDC 100 Philosophie & Psychologie	RVK CA - CK Philosophie	VZG	±0
DDC 100 Philosophie, Parapsychologie und Okkultismus, Psychologie	RVK CA - CK Philosophie	VZG	2013 ±0
DFG 108 Philosophie	RVK CA - CK Philosophie	Gerald Ja...	2020-05-27 ±0

Tree View

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

DDC Dewey Decimal Classification

Philosophie

DDC Dewey Decimal Classification

Available languages: de

Sources: JSKOS Server

available mappings

Cocoda live Demo

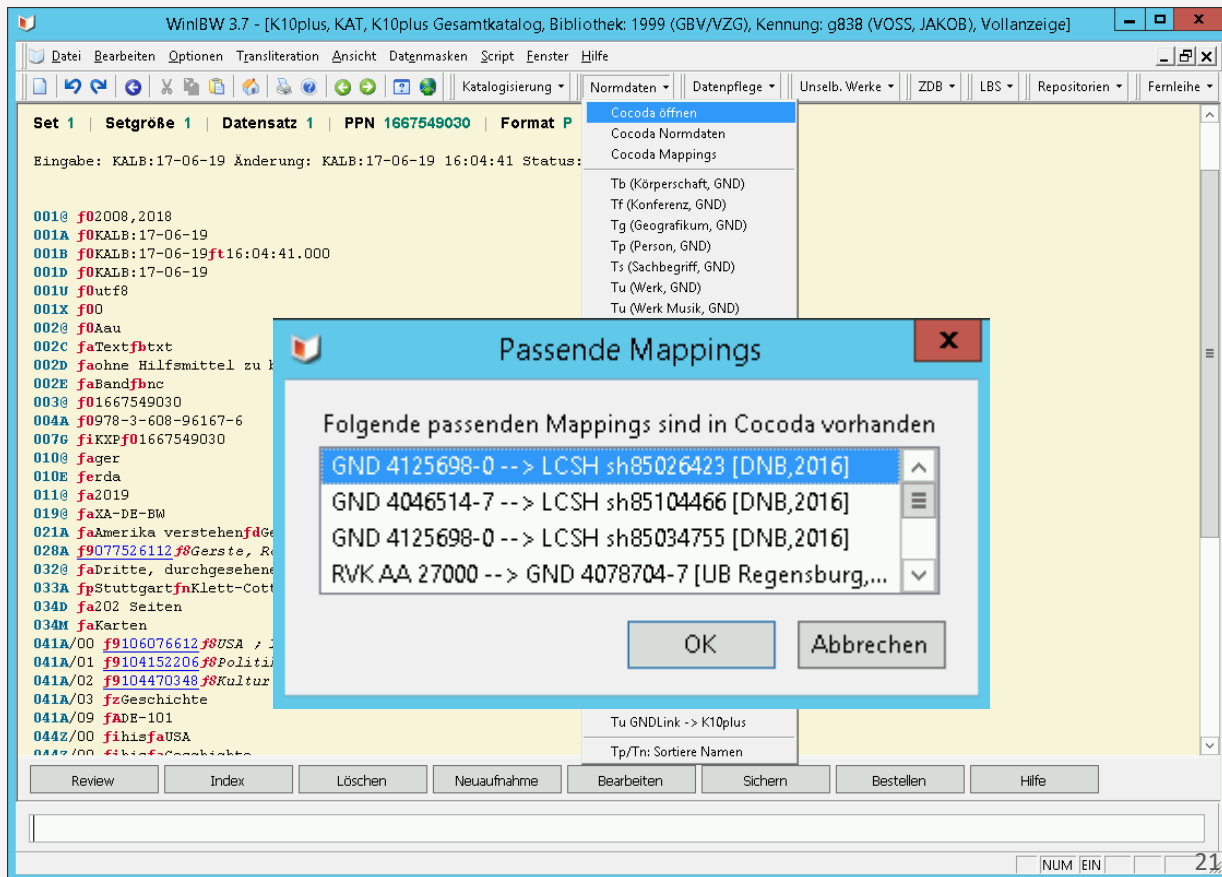
Anbindung an Portale und Anreicherung

Anbindung an externen Software

Katalogisierungstool WinIBW: Proof of Concept zum Ausprobieren

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

Sacherschließungstool Digitaler Assistent (DA-3)



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

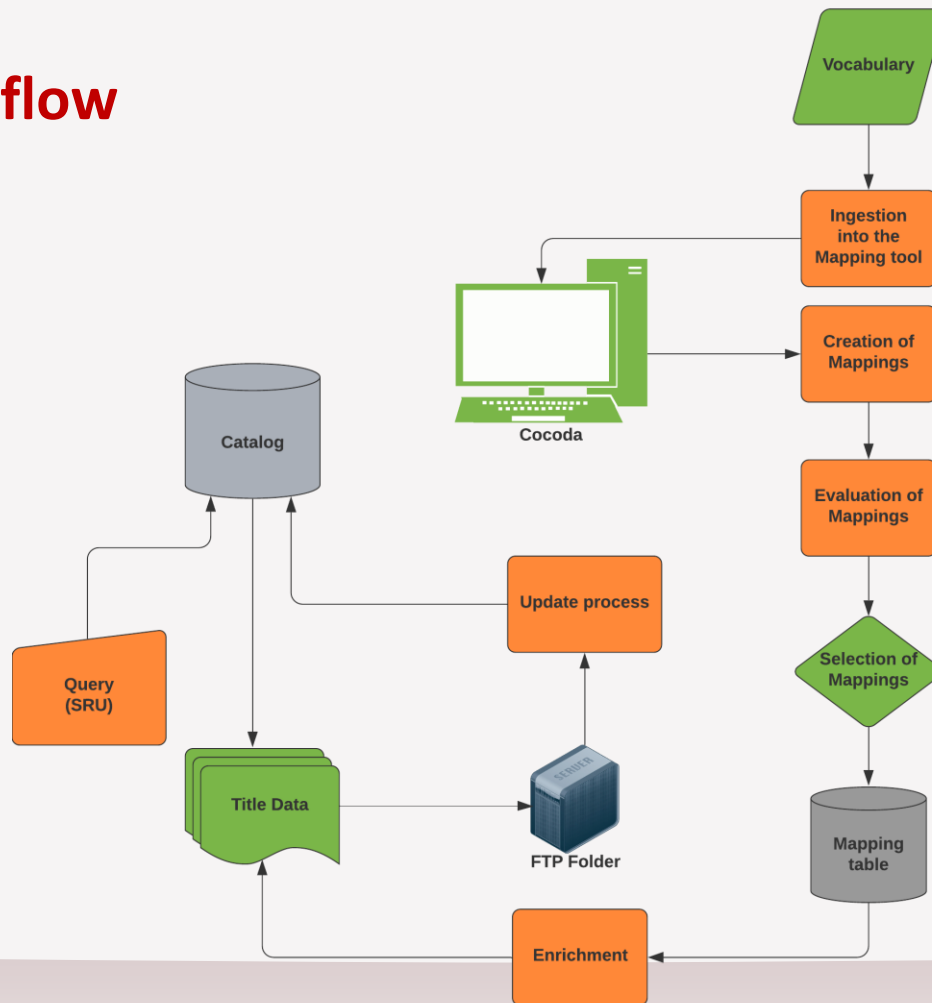
zusätzliches PICA-Feld mit Quellenangabe

\$a38.30

\$Acoli-conc DDC-RVK

\$Ahttps://coli-conc.gbv.de/api/mappings/af8ac88b-f7ab-427a-8e06-9e091d281bdc

coli-rich Workflow



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

David-Bejamin Rohrer (Systembibliothekar)
rohrer@gbv.de

