

coli-conc

Nutzung und Erstellung von Normdaten-Mappings

Jakob Voß, Uma Balakrishnan, Stefan Peters

Biblio*kongress, Leipzig

20. März 2019

coli-conc: das Projekt

coli-conc

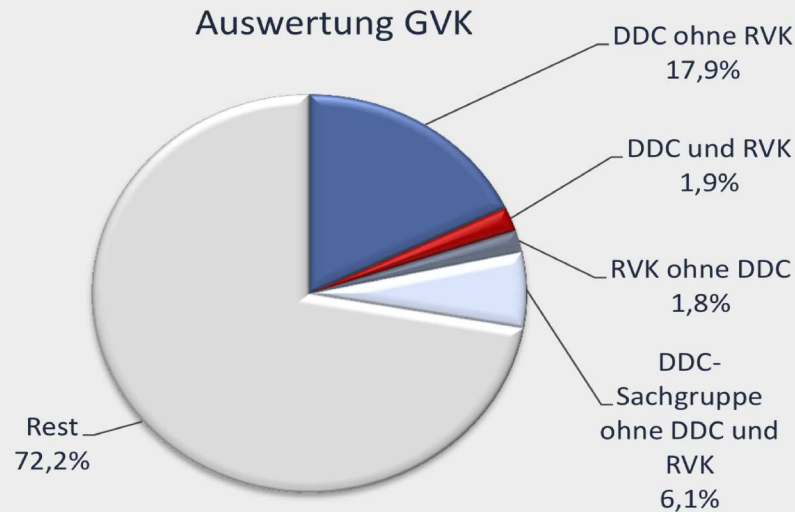
coli-conc ist ein VZG-Projekt, das Dienste für **semi-automatische Erstellung** von Mappings und Konkordanzen zwischen KOS, **Nutzung, Pflege und Speicherung** der KOS und deren Mappings in einem **einheitlichen Format** auf einer **einheitlichen Plattform** in **kooperativer Arbeit** ermöglicht.

Motivation

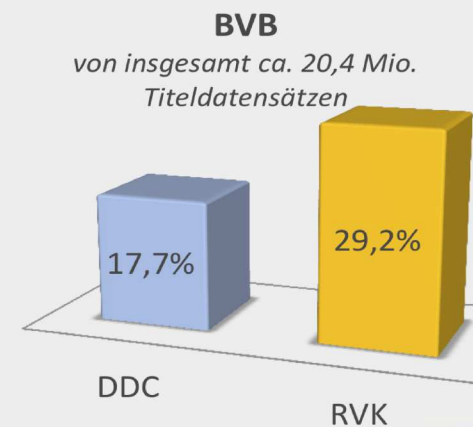
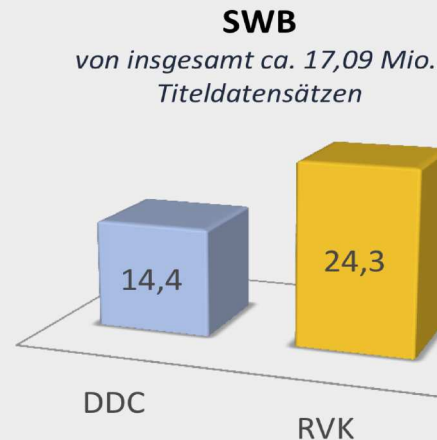
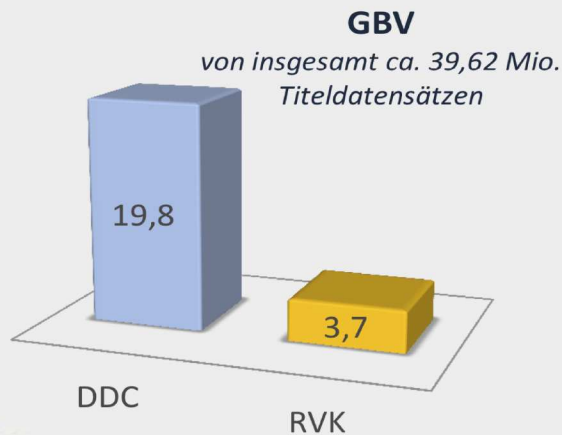
Desiderat über zwei Jahrzehnte-Zusammenführung von zwei Welten durch Bildung einer **vollständigen Konkordanz**, und zwar zwischen der **Regensburg Verbundklassifikation (860.000 Klassen)** und der **Dewey Dezimal Klassifikation (ca. 48.000 Klassen)** zu verwirklichen.

Statistik – Stand 2012

Titeldatensätze in den Verbundkatalogen mit DDC- und RVK-Notationen

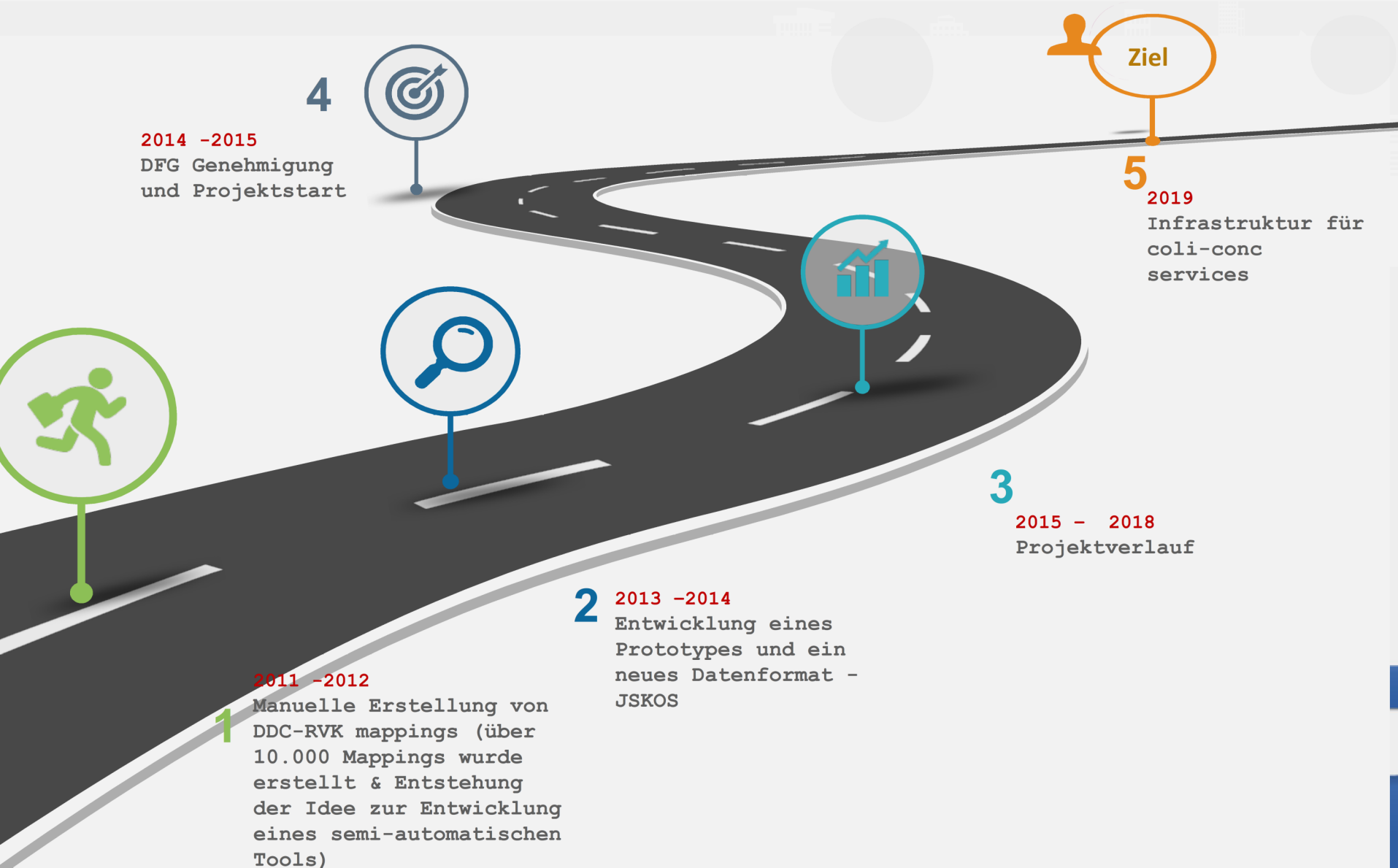


von insgesamt ca. 39,62 Mio. Titeldatensätzen

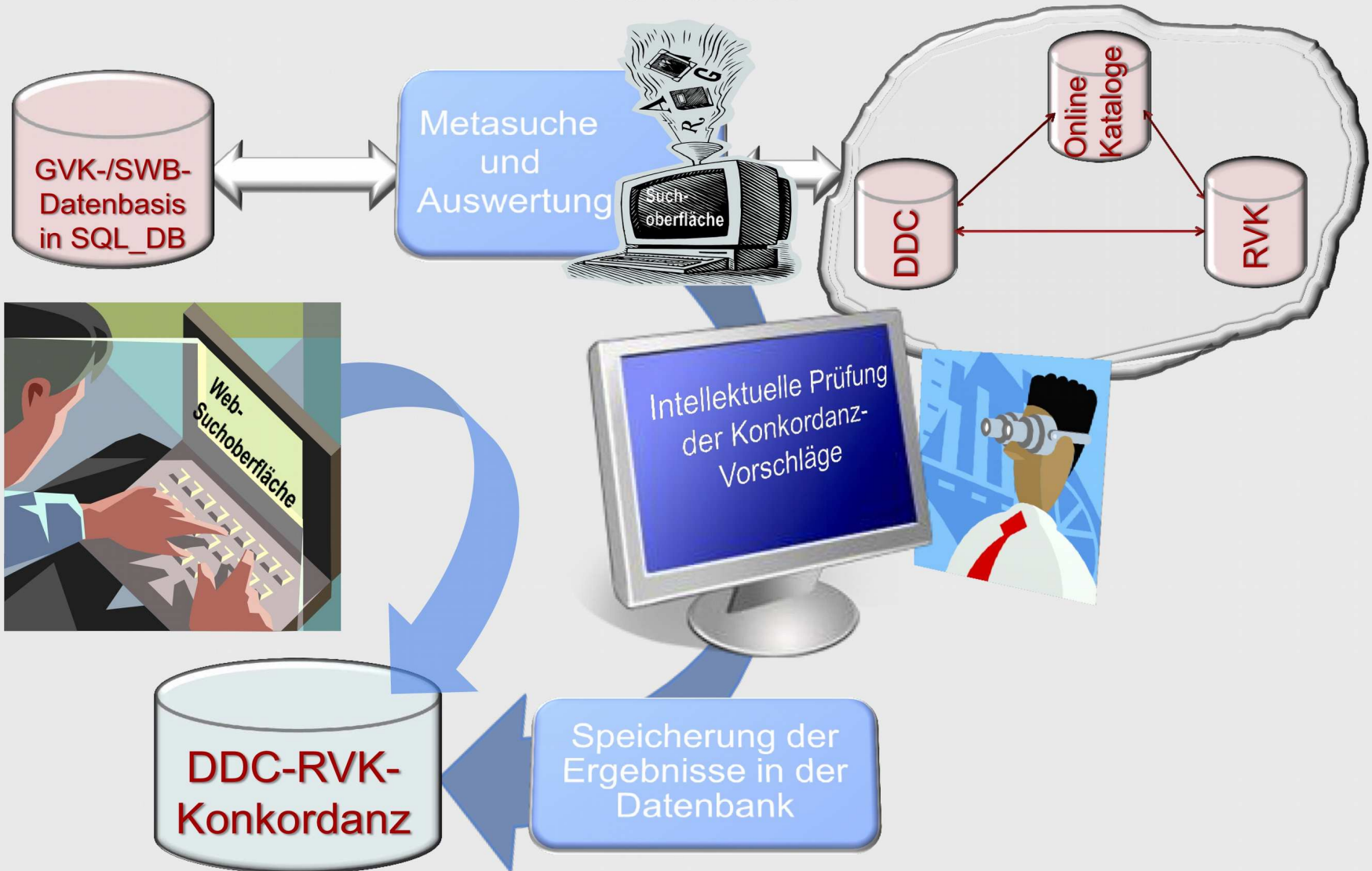


Von der Entstehung einer Idee bis zur ihrer Realisierung

The Roadmap



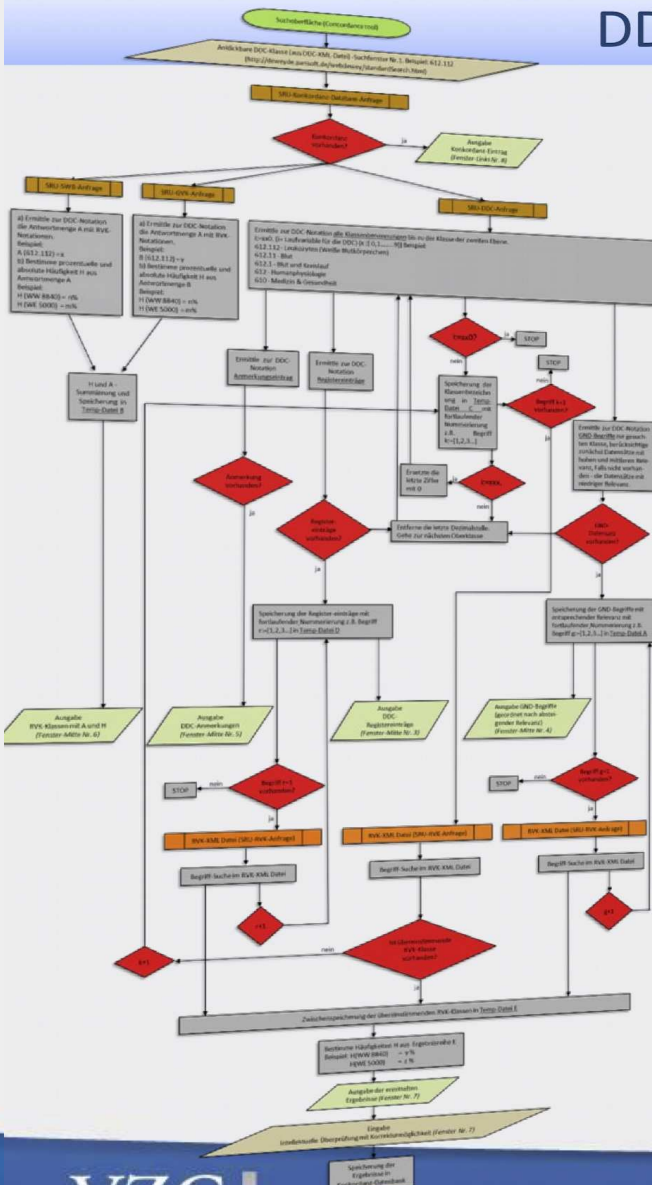
Ausblick



Mapping-Algorithmus und erster Entwurf

https://www.gbv.de/Verbundzentrale/Publikationen/publikationen-der-vzg-2013/pdf/Balakrishnan_131120_RVK_WS_Konkordanz.pdf

DDC-RVK Concordance tool (1)



DDC - RVK Concordance tool
DDC - RVK Concordance Database

DDC-System (über SRU-Schnittstelle) (Fenster Nr. 1)

600 Technik, Medizin, angewandte Wissenschaften
 610 Medizin & Gesundheit
 612 Humanphysiologie
 612.1-612.8 Einzelne Funktionen, Systeme, Organe
 612.1 Blut und Kreislauf
 612.11 Blut
 612.112 Leukozyten (Weiße Blutkörperchen)
 612.1121 Biochemie
 612.[.1122] Biophysik

Konkordanz

Konkordanz vorhanden (Fenster Nr. 8)

DDC	RVK
612.112	WW 8840 WW 8844

RVK-System (XML-Datenbank) (Fenster Nr. 2)

WW 8640 - WW 9239 Körperflüssigkeiten und deren Systeme einschließlich deren Bildungsorgane
 WW 8640 - WW 8679 Taxonomischer Schlüssel (Anwendung nur bei mit (T) versehenen Systemstellen) (Die Schlüsselzahl wird der Systemstelle zugezählt)
 WW 8720 - WW 8999 Blut und Blutbestandteile
 WW 8720 - WW 8759 Taxonomischer Schlüssel (Anwendung nur bei mit (T) versehenen Systemstellen) (Die Schlüsselzahl wird der Systemstelle zugezählt)
 WW 8800 - WW 8839 Blutplasma, Blutserum
 WW 8840 - WW 8879 Blutkörperchen (Erythrozyt, Leukozyt), Hämoglobin
 WW 8960 - WW 8999 Blutgerinnung (Thrombozyt)
 WW 9040 - WW 9079 Erythropoetisches System
 XX Tiermedizin
 YB - YC Innere Medizin

Registereintrag (Fenster Nr. 3)

Leukozyten-Humanphysiologie
 Weiße Blutkörperchen-Humanphysiologie

GND-Datensatz (Fenster Nr. 4)

Hohe Relevanz
 Leukozyt
 Mittlere Relevanz
 Alkalische Leukozytenphosphatase
 Blutlymphozyt
 Granulozyt
 Leukozytenadhäsion
 Leukozytenphosphatase
 Monozyt

Anmerkungen (Fenster Nr. 5)

Notation stillgelegt; klassifiziere in 612.112

Trefferliste SWB/GVK (Fenster Nr. 6)

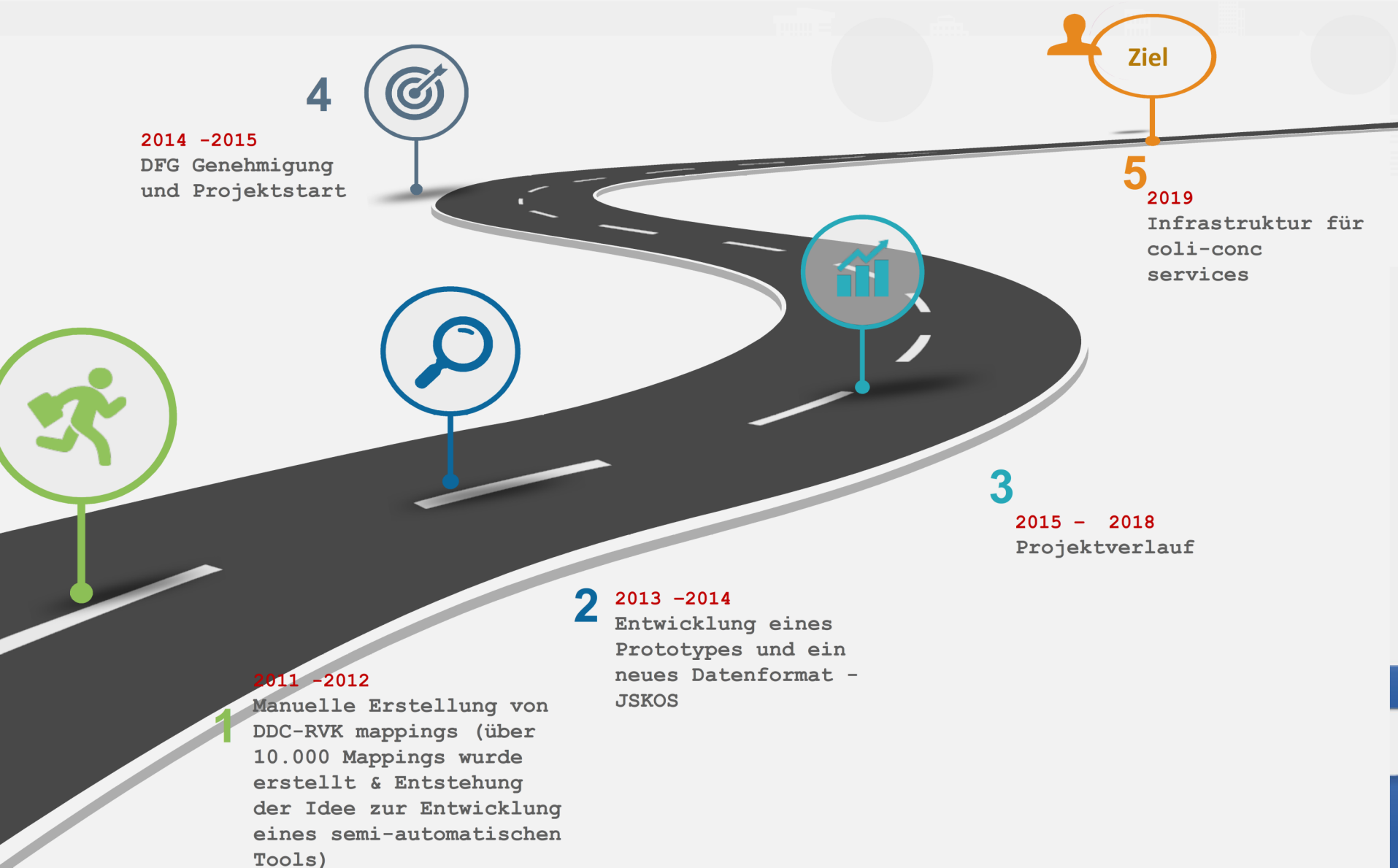
Anzahl von RVK-Notationen in der Trefferliste - 22
 WW 8840 - 68,2% (15 Treffer)
 WC 1500 - 4,5% (1 Treffer)
 WE 2400 - 4,5% (1 Treffer)
 WE 5000 - 4,5% (1 Treffer)
 WF 9870 - 4,5% (1 Treffer)
 WW 8844 - 4,5% (1 Treffer)
 WW 9124 - 4,5% (1 Treffer)
 YC 2500 - 4,5% (1 Treffer)

DDC-RVK Konkordanz (Fenster Nr. 7) (für intellektuelle Überprüfung)

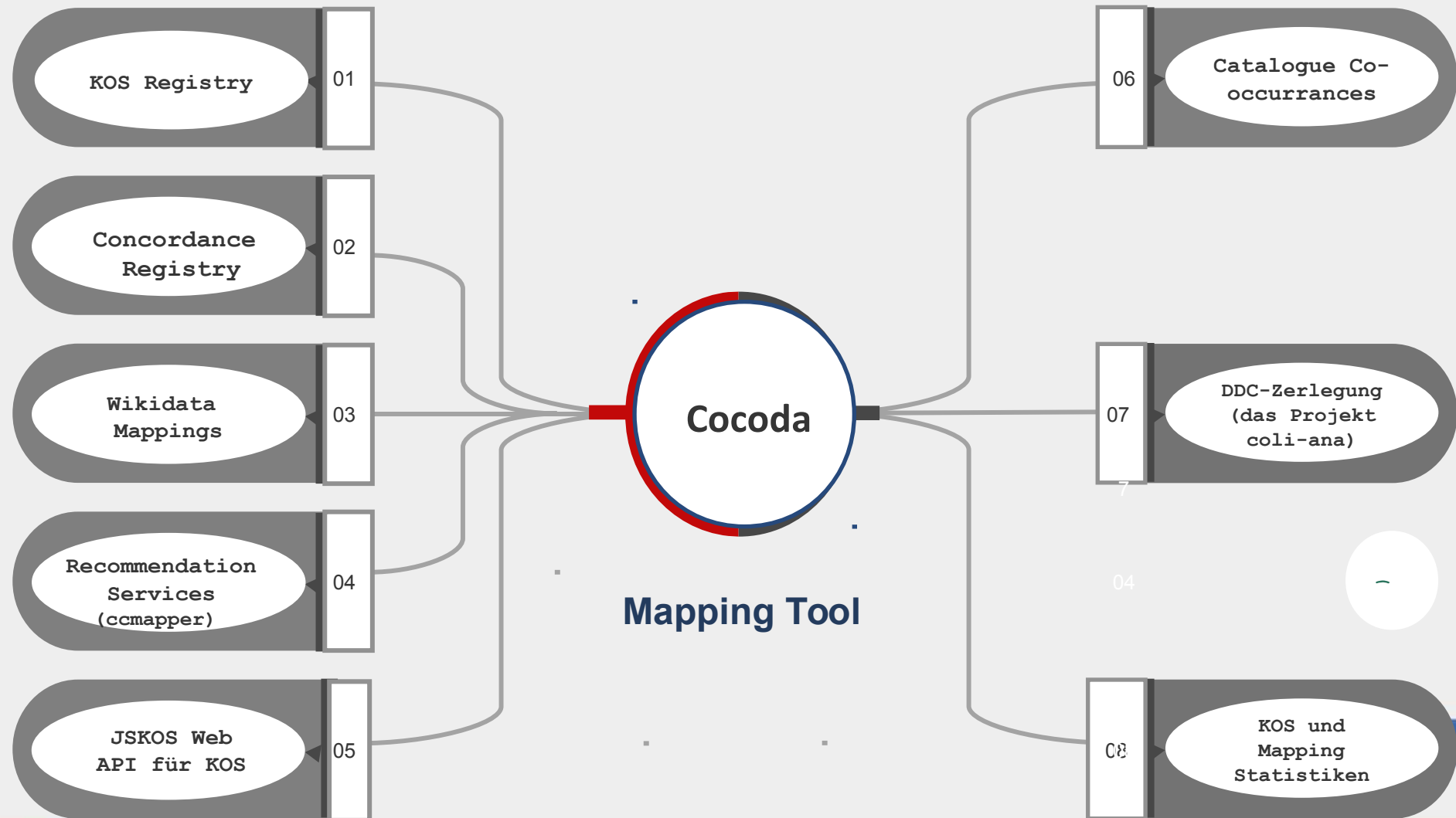
DDC	RVK
612.112	WW 8840
612.112	WW 8844
612.112	YC 2500

Von der Entstehung einer Idee bis zur ihrer Realisierung

The Roadmap



coli-conc Services



Cocoda Workflow



Cocoda: die Webanwendung

Quell-Vokabular

alle ↕

- ★ RVK Regensburger Verbundklassifikation
- ★ BK Basisklassifikation
- ★ WD Wikidata
- ★ DDC Dewey-Dezimalklassifikation
- ★ GND Gemeinsame Normdatei

[Alle 96 anzeigen](#)

Ziel-Vokabular

alle ↕

- ★ RVK Regensburger Verbundklassifikation
- ★ BK Basisklassifikation
- ★ WD Wikidata
- ★ DDC Dewey-Dezimalklassifikation
- ★ GND Gemeinsame Normdatei

[Alle 96 anzeigen](#)

Willkommen bei Cocoda!

[Anleitung](#) - [Dokumentation](#) - [GitHub](#)

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

★ BK Basisklassifikation

BK Basisklassifikation

<http://uri.gbv.de/terminology/bk/><http://bartoc.org/en/node/18785>

Created: 15. März 2019

Modified: 15. März 2019

Available languages: de

Available mappings

Top Concepts:

- ↓ 0 Allgemeine Werke und Philosophie
- ↓ 1-2 Geisteswissenschaften
- ↓ 3-4 Naturwissenschaften
- ↓ 5 Ingenieurwissenschaften
- ↓ 7-8 Sozialwissenschaften

Concept Tree +

Mapping Editor

Drag and drop a
concept here.Drag and drop a
concept here.

Mapping Browser

Available Mappings ▾



Automatic Mappings ▾



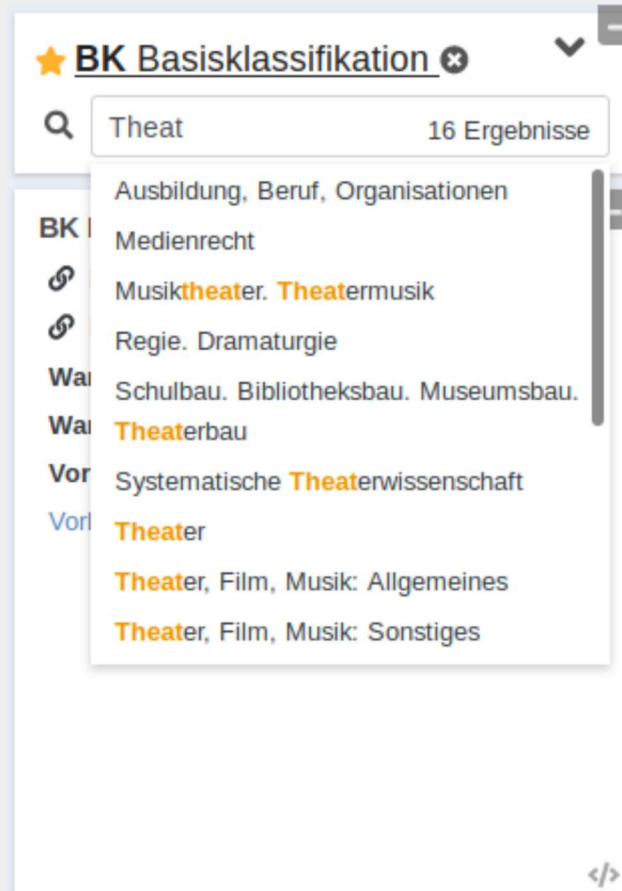
Select a concept to see mappings and mapping recommendations

Target Scheme

all ▾

[Show 96 more](#)

Auswahl von Konzepten



★ **BK Basisklassifikation** ✕

🔍 Theat 16 Ergebnisse

- Ausbildung, Beruf, Organisationen
- Medienrecht
- 🔗 Musik**theater**. **Theat**ermusik
- 🔗 Regie. Dramaturgie
- Wa Schulbau. Bibliotheksbau. Museumsbau.
- Wa **Theater**bau
- Vor Systematische **Theater**wissenschaft
- Vor **Theater**
- Vor **Theater**, Film, Musik: Allgemeines
- Vor **Theater**, Film, Musik: Sonstiges

</>

★ BK Basisklassifikation ⊕

🔍 Theater

Concept Detail ⊕

- > 0 Allgemeine Werke und Philosophie
- ▼ 1-2 Geisteswissenschaften
 - > 10 Geisteswissenschaften allgemein
 - > 11 Theologie
 - > 15 Geschichte
 - > 17 Sprach- und Literaturwissenschaft
 - > 18 Einzelne Sprachen und Literaturen
 - > 20 Kunstwissenschaften
 - > 21 Einzelne Kunstformen
 - ▼ 24 Theater ⊕
 - > 24.00 Theater, Film, Musik:
Allgemeines
 - > 3-4 Naturwissenschaften
 - > 5 Ingenieurwissenschaften
 - > 7-8 Sozialwissenschaften

Mapping Editor

Drag and drop a
concept here.Drag and drop a
concept here.

Mapping Browser

Available Mappings ▾

C L W

Automatic Mappings ▾

CR CC CO

From

To



Target Scheme

Search for schemes...

all ▾

[Show 96 more](#)

Mapping-Browser

- Lokal gespeicherte Mappings
- Zentral gespeicherte Mappings
 - Gesammelte Konkordanzen (Concordance Registry)
 - **Neu eingetragene Mappings**
- Mapping-Vorschläge
 - Abgleich von Klassenbenennungen (CR)
 - Katalog-Occurrences (CO)
 - Weitere Algorithmen (CC)
 - Reconciliation-Services (GND, Wikidata...)

Mapping-Browser



Vorhandene Mappings ▾

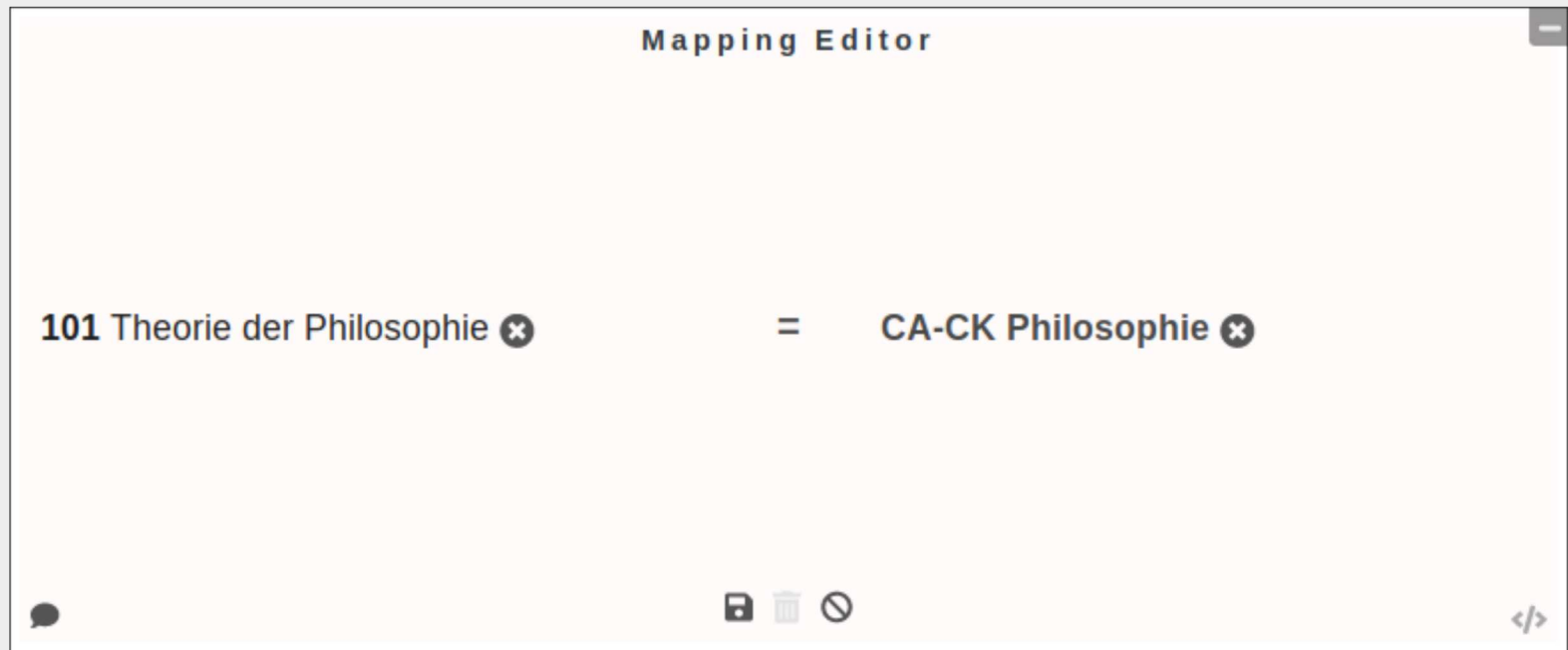
C **L** **W**

Automatische Mappings ▾

CC **CO** **WR**

	Von		Zu		Erstellt von	
C	RVK	D Pädagogik	=	BK	80 Pädagogik	Jakob Voß ±0 ⓘ ↗
C	DDC	370 Bildung und Erziehung		RVK	D Pädagogik	VZG ⓘ ↗
L	Lokal: keine Ergebnisse					
W	Wikidata-Mappings: keine Ergebnisse					
CC	RVK	D Pädagogik		DDC	370 Bildung und Erziehung	ⓘ ↗ 🔒
CC	"	D Pädagogik		"	370.12 Klassifikation und philosophische Grundlagen	ⓘ ↗ 🔒
CC	"	D Pädagogik		"	370.9 Geschichte, geografische Behandlung, Biografien	ⓘ ↗ 🔒
CC	"	D Pädagogik		"	371.26013 Validität und Reliabilität von Tests	ⓘ ↗ 🔒
CC	"	D Pädagogik		"	371.261 Testkonstruktion	ⓘ ↗ 🔒
			...			
CO	RVK	D Pädagogik	—	—		GVK - Ge... 6 ⓘ ↗
WR	RVK	D Pädagogik	≈	WD	Q7922 Pädagogik	ⓘ ↗ 🔒
WR	"	D Pädagogik	≈	"	Q19298814 Pädagogik	ⓘ ↗ 🔒
WR	"	D Pädagogik	≈	"	Q56292983 Pädagogik	ⓘ ↗ 🔒 ↗

Mapping-Editor



Accounts

Layout

Keyboard Shortcuts

Data Sources

Local Mappings

Profile

The following information is stored in all your mappings so it may also be shared:

Name

Identity (must be an URI, is automatically filled after login)

 Login via GitHub Login via ORCID Login via LDAP (VZG)

Cocoda uses a dedicated [login server](#) to authenticate users: [Impressum](#) • [Privacy Policy](#)

Mapping Provider

This provider will be used to save mappings.

Local

Mappings saved locally in the browser

<http://coli-conc.gbv.de/registry/local-mappings>

✓ Mappings

Login Server

Logged in as [Jakob Voß](#)

[Logout](#)

Name:

[Change](#)

Connected Identities



GitHub

Jakob Voß

[Disconnect](#)

<https://github.com/nichtich>



ORCID

Jakob Voß

[Disconnect](#)

<https://orcid.org/0000-0002-7613-4123>



LDAP (VZG)

[Connect](#)

[Datenschutz](#)

[Impressum](#)

[Sources](#)

[My Data](#)

[Sessions](#)

[API](#)

[Delete User Account](#)

VZG

Verbundzentrale des GBV
www.gbv.de

Concordances Search Mappings

from x to x creator x

from ↕	to ↕	description ↕	creator ↕	date ↕	download	mappings ↕
DDC	BK	Chemie	TUB Hamburg-Harburg		JSKOS CSV	99
RVK	BK	Recht	ULB Tirol	2013	JSKOS CSV	16.835
DDC	RVK	DDC 1000er-Klassen	VZG	2013	JSKOS CSV	2.050
RVK	DDC	TE-TZ Geologie und Paläontologie	GESIS		JSKOS CSV	1.163
GND	DDC	CrissCross-SWD-DDC-Mapping	DNB		JSKOS CSV	170.641
DDC	BK	Politikwissenschaft	SUB Hamburg		JSKOS CSV	655
RVK	DDC	Chemie und Pharmazie	GESIS		JSKOS CSV	771
DDC	RVK	DDC 100er-Klassen	VZG	2013	JSKOS CSV	753
DDC	RVK	Mathematik	Philipp Zumstein	2016	JSKOS CSV	40
DDC	RVK	Bibliotheks- und Informationswissenschaften	HdM Stuttgart	2009	JSKOS CSV	384
RVK	DDC	CL-CZ Psychologie	Manuela Queitsch		JSKOS CSV	455
DDC	RVK	Medizin und Gesundheit	VZG	2011	JSKOS CSV	2.813
RVK	DDC	Recht	VZG	2015	JSKOS CSV	2
DDC	BK	Full DDC up to three digits	VZG		JSKOS CSV	1.223
GND	LCSH	project macs	DNB	2016	JSKOS CSV	43.744
DDC	ixtheo	DDC zu IxTheo-Klassifikation	Timotheus Chang Whae Kim		JSKOS CSV	586
DDC	RVK	Philosophie und Psychologie (1--)	VZG		JSKOS CSV	1.070
DDC	RVK	Recht	VZG	2012	JSKOS CSV	2.247
RVK	DDC	SA-SP Mathematik	GESIS		JSKOS CSV	232
RVK	DDC	Philosophie und Psychologie	VZG	2016	JSKOS CSV	357
RVK	DDC	U Physik	GESIS		JSKOS CSV	1.336
RVK	DDC	W Biologie	GESIS		JSKOS CSV	5.349
RVK	BK	Wirtschaft	ULB Tirol	2013	JSKOS CSV	1.041
STW	GND	Ausschnitt mit 1-zu-1 Mappings	ZBW		JSKOS CSV	15.239
RVK	DDC	TA-TD Allgemeine Naturwissenschaft	GESIS		JSKOS CSV	68
RVK	GND	GND-Indexterme der RVK	UB Regensburg	2015	JSKOS CSV	94.195
RVK	BK	Germanistik	ULB Tirol	2013	JSKOS CSV	23.403

27 concordances Total: 386.772

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

VZG

Verbundzentrale des GBV
www.gbv.de

Weitere Werkzeuge

Excel

- Daten sind einfach zu bearbeiten
- Sehr fehleranfälliges Format (Validierung?)
- Wir unterstützen CSV-Dateien (begrenzt)

ccmapper (pansoft)

- Mapping-Tool für **RVK → DDC**
- Weiterentwicklung von ccmapper-no
 - Humord & Realfagstermer → DDC
- Algorithmen für Mapping-Vorschläge
- Eigenes Interface
- Wir können Accounts zum Testen vergeben!

[Mapping](#) [Review](#) [Suche](#)

RVK

V Chemie und Pharmazie > VH Anorganische Chemie > VH 5000 - VH 9759 Anorganische Chemie allgemein

VH 5000 - VH 9759
Anorganische Chemie allgemein

Nächstes

Aktivitäten

 jv at 04. Mär 2019 12:42
 jv at 04. Mär 2019 12:45

Andere

Status

 Review erledigt

Suche nach dem gewählten Ausdruck

WebDewey DE	Wikipedia DE
WebDeweySearch	Wikipedia EN
Cocoda Master	Cocoda Dev
RVK Online	

Anorganische Chemie allgemein

Suche

GND Index

[Seite nach oben](#) [Seite nach unten](#)

Anorganische Beschichtung	667.9
Anorganische Chemie allgemein , GND Eintrag nicht gefunden	
Anorganische Chemie	546
Anorganische Peroxide	546.7212
Anorganische Polymere	541.2254

Mapping

Mapping Kandidaten (9)


Suche WebDewey

-
- Konzept Mapping erledigt
-
-
- Konzept Mappings exportieren

Speichern

Sortierreihenfolge Mappings mit Kommentaren


615.92 Anorganische Gifte

D (4.026)    jv at 04. Mär 2019 12:42 : geändert von OFFEN nach ~RM

553.9 Anorganische Gase

D (3.823)    jv at 04. Mär 2019 12:42: geändert von OFFEN nach ≈CM

+ 572.514 Anorganische Substanzen--Stoffwechsel, . . .

D (3.580)    jv at 04. Mär 2019 12:42: geändert von OFFEN nach RJ

615.2 Anorganische Arzneistoffe

D (3.580)  

Dewey

Technik > Medizin & Gesundheit > Pharmakologie und Therapeutik > Toxikologie

615.92 Anorganische Gifte

Für gasförmige anorganische Gifte siehe 615.91
Für Vergiftungen durch Strahlung siehe 616.9897

Registereintrag

Anorganische Gifte--Humantoxikologie

Geerbte Anweisungen (2)

Untergeordnete Klassen (3)

615.9 Toxikologie

Hier auch: Gifte und Vergiftungen
Klassifiziere forensische Toxikologie in 614.13
Klassifiziere Wirkungen von Giften auf ein bestimmtes System oder Organ bei dem System oder Organ in 616-618 , plus Notation 616.1-616.9:071 aus der Anhängetafel bei 616.1-616.9 , Notation 617:071 aus der Anhängetafel bei 617 , oder Notation 618.1-618.8:071 aus der Anhängetafel bei 618.1-618.8 , z.B. Wirkung von Giften auf die Leber 616.362071

615 Pharmakologie und Therapeutik

Klassifiziere umfassende Werke über Pharmakologie in 615.1
Klassifiziere umfassende Werke über Therapeutik in 615.5

Digitaler Assistent (DA-3)

- Werkzeug zur Unterstützung der Sacherschließung
- Verwendet bereits Konkordanzen
- Anbindung an Cocoda angedacht

Ausblick

Nächste Schritte

- Mapping-Übersichten und Statistiken
- Usability-Tests und Verbesserung
- Anreicherung von RVK-BK-Mappings in K10plus
- Integration weiterer KOS (u.A. Museumsvokabulare)
- coli-ana (DDC-Zerlegung) und DDC-Number-Building
- Regeln und Workflows für Qualitätsicherung
- Anbindung an Digitalen Assistenten u.A. externe Tools

Probieren Sie Cocoda aus!

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

Kurzeinführung im Screencast

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

*Erster coli-conc Workshop
voraussichtlich am 11. Juni in Göttingen*