

coli-conc

DFG Projektantrag

**im Bereich
Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS)
Förderprogramm Werkzeuge und Verfahren des wissenschaftlichen
Informationsmanagements**

Reiner Diedrichs (VZG)*

27. November 2014**

* *Projektkoordination: U. Balakrishnan (balakrishnan@gbv.de)
Technische Koordination: J. Voß (voss@gbv.de)*

** *Gekürzte Fassung veröffentlicht am 18. August 2015
(DOI [10.5281/zenodo.28290](https://doi.org/10.5281/zenodo.28290))*

Inhaltsverzeichnis

I. Beschreibung des Vorhabens.....	1
1. Ausgangslage und eigene Vorarbeiten.....	2
1.1 Projektbezogene Publikationen.....	5
1.1.1 Veröffentlichte Arbeiten aus Publikationsorganen mit wissenschaftlicher Qualitätssicherung, Buchveröffentlichungen sowie bereits zur Veröffentlichung angenommene, aber noch nicht veröffentlichte Arbeiten.....	5
1.1.2 Andere Veröffentlichungen.....	5
2 Ziele und Arbeitsprogramm.....	6
2.1 Voraussichtliche Gesamtdauer des Projekts.....	6
2.2 Ziele.....	6
3 Literaturverzeichnis.....	8
II. Anlangen.....	I
A. Vorhandene DDC- und RVK-Konkordanzen.....	I
B. Algorithmus zur Erstellung von DDC-RVK Konkordanzen.....	III
C. Vorhandener Prototyp der Benutzeroberfläche.....	V
D. Systemübersicht der geplanten Infrastruktur.....	VI
H. Letters of Intent.....	VII

I. Beschreibung des Vorhabens

Entwicklung einer Infrastruktur für den Austausch, die Erstellung, und die Wartung von Konkordanzen zwischen bibliothekarischen Wissensorganisationssystemen.

Die Nutzung und der Austausch von Wissensorganisationssystemen (KOS) hat in den letzten Jahren durch die Verbreitung geeigneter Standards und Werkzeuge deutlich zugenommen. Während diese KOS vermehrt zur strukturierten Erschließung von Dokumenten und Daten angewandt werden, sind Konkordanzen zwischen verschiedenen Systemen nur sehr begrenzt verfügbar. Mit dem im vorliegenden Antrag beschriebenen Projekt sollen Werkzeuge und Verfahren geschaffen werden, um sowohl den Austausch und die Nutzung, als auch die intellektuelle Erstellung und Qualitätsbewertung von bibliothekarischen Konkordanzen zu vereinfachen und zu beschleunigen. Im Projekt sollen wiederverwendbare Modulen zum einheitlichen Zugriff auf Wissensorganisationssysteme, Konkordanzen und Konkordanzbewertungen erstellt werden, die zusammen eine Webanwendung zur effektiven Bearbeitung von Konkordanzen bilden. Dabei werden vorhandene Programme (KOS-Software, -Portale, Mapping-Algorithmen, Katalogisierungssoftware etc.) evaluiert und gegebenenfalls erweitert sowie neue Komponenten für die Speicherung, Bereitstellung und Analyse von verschiedenen Konkordanzen erstellt. Alle Komponenten sollen so über einheitliche und offene APIs miteinander verbunden werden, so dass eine verteilte Infrastruktur für Erstellung, Austausch und Wartung von Konkordanzen entsteht.

Development of an infrastructure to facilitate management and exchange of concordances between library knowledge organization systems

The use and exchange of knowledge organization systems (KOS) has increased due to the availability of tools and standards during the last few years. These KOS play a vital role in the management of knowledge and data. However, concordances between these systems, are rather rare. The proposed project aims to address this gap by developing tools, methods and techniques to simplify and accelerate both the intellectual creation of concordances as well as ease the use and exchange of the same and at the same time provide quality monitoring that aid quality management. The tools will be provided as modular web applications to support effective processing of concordances. In addition, existing software (KOS software, mapping algorithms, cataloguing software ...) will be evaluated and enhanced with new components for storage, access to and analysis of different concordances. These components will be linked with each other through uniform and open APIs, so that a shared infrastructure is developed for the management, exchange and building of concordances.

1. Ausgangslage und eigene Vorarbeiten

Ausgangslage

Die Standardisierung des Simple Knowledge Organization System (SKOS) hat im Rahmen von Linked Open Data (LOD) dazu beigetragen, dass Wissensorganisationssysteme (KOS) wie Klassifikationen, Normdateien und Thesauri zunehmend frei zur Verfügung gestellt werden (Hanke 2014). Zudem wurden in den letzten Jahren mehrere Programme zur Verwaltung und Bereitstellung von KOS entwickelt (iQvoc,¹ Poolparty,² VocBench,³ OpenSKOS,⁴ Skosmos,⁵ xTree⁶ etc.) sowie mehrere Portale zur Sammlung und Publikation vorhandener KOS aufgebaut (OCLC Terminology Services,⁷ Museumsvokabular.de,⁸ BARTOC⁹ etc.; siehe auch Golub 2009). Die einzelnen KOS oder Teile davon werden dabei idealerweise zum Download und über eigene APIs bereitgestellt.

Anders ist die Situation bei Konkordanzen (auch bekannt als Cross-Konkordanzen oder Terminologie-Mappings), die Begriffe unterschiedlicher KOS aufeinander abbilden (Mayr 2010; Keil 2012). Konkordanzen dienen der Verbesserung des Information Retrieval, vor allem in verteilten und mehrsprachigen Datenquellen. Auch zur einheitlichen Erschließung mit mehreren KOS sind Konkordanzen schon länger gewünscht (Junger 1998). Die intellektuelle Erstellung von Konkordanzen ist jedoch sehr aufwändig und stößt angesichts der Tiefe, Größe und kontrastiven Natur verschiedener KOS insbesondere bei fein gegliederten Klassifikationssystemen wie Dewey Dezimalklassifikation (DDC, etwa 38.000 Klassen) und Regensburger Verbundklassifikation (RVK, etwa 850.000 Klassen) an praktische Grenzen. Gleichzeitig wird eine vollständige Konkordanz speziell zwischen diesen Systemen schon länger gefordert (Hermes 1998): die DDC wurde im Jahr 2000 von der DNB eingeführt und auch von anderen europäischen Nationalbibliotheken eingesetzt, um Einheitlichkeit in der Klassifikationslandschaft zu schaffen und den Zugang zu und Austausch der Daten auf internationaler Ebene zu erleichtern (Heiner-Freiling 2000) und die RVK ist neben der eher begrenzten Basisklassifikation (BK) das am meisten verbreitete Klassifikationssystem im deutschsprachigen Raum.

Die genannten Schwierigkeiten bei der Konkordanzerstellung führen dazu, dass wenige vollständige Konkordanzen existieren (Balakrishnan 2013). Zu den vorhandenen Projekten, in denen in größerem Umfang Konkordanzen erstellt wurden, zählen MACS¹⁰ (seit 1997), Renardus¹¹ (1999-2002), CrissCross¹² (2006-2010) und KoMoHe¹³ (2004-2007, siehe Mayr 2008). Eine vollständige Konkordanz zwischen den großen im deutschsprachigen Raum genutzten bibliothekarischen Klassifikationen DDC und RVK existiert bislang jedoch nicht (Anlage A. *Vorhandene DDC- und RVK-Konkordanzen*). Automatische Verfahren zur

1 <http://iqvoc.net/>

2 <http://www.poolparty.biz/>

3 <http://vocbench.uniroma2.it/>

4 <http://openskos.org/>

5 <https://github.com/NatLibFi/Skosmos>

6 <http://xtree.digicult-verbund.de/>

7 <https://info.gbv.de/www.oclc.org/research/activities/termservices.html>

8 <http://www.museumsvokabular.de/>

9 <http://www.bartoc.org/>

10 http://www.dnb.de/DE/Wir/Kooperation/MACS/mac_node.html

11 <http://www.ukoln.ac.uk/metadata/renardus/>

12 <http://www.dnb.de/DE/Wir/Projekte/Abgeschlossen/crisscross.html>

13 <http://www.gesis.org/en/research/external-funding-projects/archive/komohe/>

Erstellung von Konkordanzen sind zwar möglich und Gegenstand der Forschung (siehe OAEI Library Track, 2012-2014 und z.B. Pfeffer 2013), diese Verfahren sind jedoch eher für Mapping-Vorschläge geeignet und in der Regel unvollständig (Lauser 2008; Reiner 2010). Vorhandene Konkordanzen decken daher oft lediglich Teile der von ihnen gemappten KOS ab.

Hinzu kommt, dass die Ergebnisse von Konkordanz-Projekten meist nur intern genutzt werden und zum Teil nicht öffentlich zugänglich sind (Balakrishnan, 2013). Falls Konkordanzen doch weitergegeben werden, beispielsweise über persönliche Kontakte, geschieht dies eher als Teil von anderen Datenbeständen oder in nicht-standardisierten Formaten ohne genauere Angaben über Methodik und Aktualität. Zu den wenigen Ausnahmen frei zugänglicher Konkordanzen gehören das Projekt VIAF von OCLC (OCLC 2014),¹⁴ eine Sammlung von Konkordanzen der ZBW¹⁵ und das Projekt Wikidata¹⁶ der Wikimedia Foundation (Voß 2014, Vrandečić 2014). Das Potential offener Community-Projekte wie Wikidata für die Erstellung von Konkordanzen ist zwar groß (vgl. Hanke 2014 und Soergel 2011) und die in diesen Projekten erarbeiteten Folksonomies¹⁷ lassen sich auf etablierte KOS abbilden (Voß 2007; 2012) – offen bleibt jedoch auch hier die Qualität der so erstellten Konkordanzen.

Die geringe Verfügbarkeit von Konkordanzen hat zur Folge, dass sich bisher, trotz Bemühungen zur Standardisierung (siehe SKOS, ISO 25964-2, Keil 2012, Soergel 2011), weder Austauschformate noch Verfahren für die nachhaltige Pflege und Bereitstellung durchgesetzt haben. Vor allem im Bereich bibliothekarischer Klassifikationen sind weder Werkzeuge noch etablierte Verfahren bekannt. Auch fehlte bislang eine Übersicht existierender Konkordanzen und Konkordanz-Projekte. Grundsätzlich lässt sich feststellen, dass es sowohl an einer Infrastruktur zur Bereitstellung und zum Austausch von Konkordanzen, als auch an Werkzeugen für ihre Erstellung und qualitative Bewertung mangelt.

Vorarbeiten

Seit 2003 werden an der Verbundzentrale des GBV (VZG) im Projekt Colibri/DDC neue (semi-)automatische Verfahren für die DDC erforscht und entwickelt (Reiner 2003; 2008). Mittlerweile besteht das Projekt aus vier Teilprojekten:

- *coli-auto* seit 2003: Automatische DDC-Klassifikation (Reiner 2009; 2010)
- *coli-ana* seit 2004: Analyse von synthetischen DDC-Notationen (Reiner 2007)
- *coli-conc*, seit 2007: Erstellung von Konkordanzen von und zu DDC (Balakrishnan 2011a)
- *coli-corr*, seit 2014: Überprüfung der Korrektheit von DDC-Notationen

Als Ergebnis des Teilprojekts *coli-conc* liegen derzeit folgende Konkordanzen vor:

- Vollständige Konkordanz zwischen DDC und BK sowie zwischen DDC und RVK für Tausend Klassen der DDC bis zur dritten Ebene
- Vollständige bidirektionale Konkordanz zwischen EZB und DDC (Balakrishnan 2011b)
- Vollständige Konkordanz zwischen DDC und RVK für die Bereiche Medizin & Gesundheit und Recht (Balakrishnan 2011a; 2011b)

14 <http://viaf.org/>

15 <http://zbw.eu/stw/versions/latest/mapping/>

16 <http://wikidata.org/>

17 Folksonomies sind meist freie, von Nutzern erstellte Schlagwortsysteme.

- Konkordanz zwischen DDC und RVK für den Bereich Philosophie (in Arbeit)

Insgesamt wurden etwa 7000 Mappings zwischen DDC und RVK erstellt, die rund 10 % der vorhandenen DDC-Klassen abdecken. Neben der eigenen Erstellung von intellektuell geprüften Konkordanzen wurde 2011 eine Umfrage durchgeführt, um den Stand von Konkordanzen mit Schwerpunkt auf DDC und RVK im deutschsprachigen Raum zu ermitteln (siehe Balakrishnan 2013 und Anlage A. *Vorhandene DDC- und RVK-Konkordanzen*). Die Ergebnisse von Umfrage und Konkordanzerstellung (Art der Konkordanzbildung, anwendbare Methoden, Erstellungsschwierigkeit...) wurden analysiert, um sie für die Erstellung weiterer Konkordanzen zu nutzbar zu machen (Balakrishnan 2011a; 2011b).

Neben der Arbeit an bibliothekarischen Wissensorganisationssystemen wurden an der VZG von Communities erstellte Tagging-Systeme und Folksonomies untersucht (Voß 2007; 2012; 2014). Insbesondere gibt es Vorarbeiten zur Extraktion von Mappings zwischen verschiedenen KOS aus der kollaborativen Datenbank Wikidata (Voß 2014; Krötzsch 2014). Eine erste Analyse zeigt, dass Wikidata Mappings mit über 300 Normdateien, Klassifikationen und andere KOS enthält und somit zur Erstellung und Verbreitung von Konkordanzen herangezogen werden sollte. Die Abdeckung und Qualität ist jedoch sehr uneinheitlich und bedarf einer genaueren Analyse.

Zur Vorbereitung des in diesem Antrag beschriebenen Projekt werden derzeit mit begrenzten Mitteln folgende Maßnahmen durchgeführt:

- Sammlung und Überführung vorhandener Konkordanzen in ein einheitliches Format, um sie in einer Datenbank für die gemeinsame Nutzung zusammenzuführen (vgl. AP 3).
- Umsetzung eines Mapping-Algorithmus basierend auf den im Projekt *coli-conc* intellektuell durchgeführten Verfahren (vgl. AP 4 und Anlage B. *Algorithmus zur Erstellung von DDC-RVK Konkordanzen*).
- Prototyp einer Benutzeroberfläche (vgl. AP 2).

Ein funktionsfähiger Prototyp der Benutzeroberfläche setzt bereits den Zugriff auf RVK und GND über offene APIs um. Die Anbindung der GND über die API von lobid.org und RVK über die API der UB Regensburg führte bereits zu einem produktiven Austausch mit den Betreibern dieser KOS und einigen Verbesserungen an ihren Diensten – diese Entwicklung deutet bereits die positiven Effekte des beantragten Projektes für vorhandene KOS an. Über die Benutzeroberfläche können unterschiedliche Begriffe der verschiedenen KOS gesucht und mit einander verknüpft werden. Teile des Programms wurden bereits als unabhängige AngularJS-Module¹⁸ veröffentlicht (Voß & Horn 2014). Ein Screenshot des Prototyps befindet sich in Anlage C. *Vorhandener Prototyp der Benutzeroberfläche*.

¹⁸ AngularJS ist ein von Google entwickeltes JavaScript-Framework zur Entwicklung von modularen Webanwendungen (siehe <https://angularjs.org/>).

1.1 Projektbezogene Publikationen

Die folgenden Publikationen der Projektmitarbeiter beinhalten Vorarbeiten zum Projekt.

1.1.1 Veröffentlichte Arbeiten aus Publikationsorganen mit wissenschaftlicher Qualitätssicherung, Buchveröffentlichungen sowie bereits zur Veröffentlichung angenommene, aber noch nicht veröffentlichte Arbeiten

- Balakrishnan, U. (2011a): Challenges in mapping classification systems – lessons from building a DDC-RVK concordance. Master Thesis.
- Reiner, U. (2007): Automatic Analysis of Dewey Decimal Classification Notations. In: *Data Analysis, Machine Learning and Applications. Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V.*, S. 697-704. DOI 10.1007/978-3-540-78246-9_82
<http://archiv.tu-chemnitz.de/pub/2007/0139/data/reiner.pdf>
- Voß, J. & Horn, M. (2014): Exposing library services with AngularJS. In: *Code4Lib Journal* 26. <http://journal.code4lib.org/articles/10023>
- Voß, J. (2012): Linking Folksonomies to Knowledge Organization Systems. In: *Metadata and Semantics Research (MTSR 2012)*, S. 89-97. DOI 10.1007/978-3-642-35233-1_9
- Voß, J. (2007) Tagging, Folksonomy & Co - Renaissance of Manual Indexing?, 243-254. In: *10th International Symposium for Information Science (ISI 2007)*.
<http://arxiv.org/abs/cs.IR/070107>

1.1.2 Andere Veröffentlichungen

- Balakrishnan, U. (2013): Das Projekt coli-conc: Ein Bericht zur semi-automatischen Erstellung von Konkordanzen zur Dewey Dezimalklassifikation. In: *VZG Aktuell*, Ausgabe 1/2013, S. 12-16.
- Balakrishnan, U. (2013): Das VZG-Projekt “coli-conc”- Brückenbildung zwischen DDC und RVK
- Voß, J. et al. (2014): Normdaten in Wikidata. lulu. ISBN 978-1-291-85658-3
- Balakrishnan, U. (2011b): An EZB-DDC Concordance. VZG-Colibri-Bericht 1/2011, VZG. <http://dx.doi.org/10.5281/zenodo.28263>

2 Ziele und Arbeitsprogramm

2.1 Voraussichtliche Gesamtdauer des Projekts

24 Monate.

2.2 Ziele

Der vorliegende Projektantrag verfolgt drei zusammenhängende Ziele:

1. Effektivere Erstellung und Pflege von Konkordanzen zwischen bibliothekarischen KOS
2. Verbesserung der Qualität von Konkordanzen
3. Förderung des Austauschs und der Nutzung von Konkordanzen und Wissensorganisationssystemen

Ziel 1

Die Erstellung von Konkordanzen ist sowohl intellektuell als auch in der praktischen Umsetzung sehr aufwändig. Vor allem zur kooperativen Bearbeitung sind bisher keine Werkzeuge bekannt. Mit der Entwicklung einer Webanwendung zur Erstellung und Bearbeitung von Konkordanzen soll diese Tätigkeit vereinfacht und für einen erweiterten Personenkreis zugänglich gemacht werden. Für die bibliothekarische Katalogisierung soll die Anwendung außerdem als tägliches Arbeitswerkzeug eingebunden werden. Automatische Verfahren zur Ermittlung von Mapping-Vorschlägen, wie sie bereits im Projekt *coli-conc* und an anderer Stelle entwickelt wurden (vgl. OAEI Library Track, 2012-2014), werden dabei zur Unterstützung der Konkordanzerstellung eingebunden. Der Schwerpunkt der Konkordanzerstellung liegt auf den im deutschsprachigen Raum weit verbreiteten Klassifikationssystemen (DDC, RVK, BK).

Maßnahmen:

- Entwicklung eines Konkordanz-Tools (AP 2)
- Bereitstellung von Konkordanzen und Wissensorganisationssystemen (AP 3)
- Weiterentwicklung und Umsetzung von automatischen Konkordanz-Verfahren (AP 4)

Ziel 2

Die beschränkte Verfügbarkeit und Nutzung von Konkordanzen und der Mangel an Standards für ihre Erstellung und ihren Austausch haben zur Folge, dass sich derzeit die Qualität von Konkordanzen nur schwer beurteilen lässt. Zur Verbesserung der Qualität müssen deshalb zunächst Werkzeuge geschaffen werden, mit denen beurteilt werden kann, ob und an welchen Stellen Ausbau, Aktualisierungen und Korrekturen von Konkordanzen sinnvoll sind. Eine qualifizierte Beurteilung erfordert zudem die möglichst breite Einbindung von Fachleuten und Anwendern.

Maßnahmen:

- Entwicklung und Umsetzung von Maßnahmen zur Qualitätsbewertung (AP 5)

- Einbindung von breiteren Nutzergruppen (AP 6)

Ziel 3

Die konsequente Veröffentlichung von Konkordanzen ist bislang nicht üblich; stattdessen werden sie meist nur intern genutzt, kaum dokumentiert und mit beschränkten Mitteln (z.B. Excel-Tabellen) verwaltet. Zur Förderung des Austauschs und der Nutzung von Konkordanzen und bibliothekarischen KOS sollen deshalb Standards geschaffen, dokumentiert und angewandt werden, mit denen sich Konkordanzen besser beschreiben, finden und nachnutzen lassen. Um der verteilten Erstellung und Nutzung von KOS gerecht zu werden, soll auch die Bereitstellung in Form einer dezentralen Infrastruktur erfolgen, in der die VZG einen Zugangspunkt betreibt.

Maßnahmen:

- Sammlung vorhandener Konkordanzen und Wissensorganisationssysteme (AP 1)
- Bereitstellung von Konkordanzen und Wissensorganisationssystemen (AP 3)
- Dokumentation und Öffentlichkeitsarbeit (AP 6)

3 Literaturverzeichnis

Eine aktualisierte Literatursammlung befindet sich verlinkt auf der Projekthomepage unter <http://coli-conc.gbv.de/publications/>.

- Golub, Koraljka & Tudhope, Douglas (2009): Terminology Registry Scoping Study (TRSS). Final Report.
- Graf, Nicole (2013): Crowdsourcing beim Swissair-Fotoarchiv. Eidgenössische Technische Hochschule Zürich, ETH-Bibliothek. <http://dx.doi.org/10.3929/ethz-a-009774132>
- Hanke, Mirko (2014): Bibliothekarische Klassifikationssysteme im semantischen Web: Zu Chancen und Problemen von Linked-Data-Repräsentationen ausgewählter Klassifikationssysteme. In: *Perspektive Bibliothek*, Band 3, Nr. 2, S. 91-119 <http://dx.doi.org/10.11588/pb.2014.2.16808>
- Hermes (1998): Die Chancen der Konkordanzen. In: *Klassifikation für wissenschaftliche Bibliotheken: Analysen, Empfehlungen, Modelle*. Deutsches Bibliotheksinstitut. S. 87-91
- Heiner-Freiling, Magdha (2000): Einführung und Nutzung der Dewey Decimal Classification (DDC) im deutschen Sprachraum. DNB
- Koch, Traugott; Neuroth, Heike; Day, Michael (2001): DDC mapping report. Renardus projekt D7.4. <http://renardus.sub.uni-goettingen.de/wp7/d7.4/>
- ISO (2013): ISO 25964-2: Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies.
- Junger, Ulrike (1998): Überlegungen zur Weiterentwicklung der Sacherschließung im Gemeinsamen Bibliotheksverbund. In: *Bibliothek - Forschung & Praxis* 22, Nr. 2, S. 252-255
- Keil, Stefan (2012): Terminologie Mapping: Grundlagen und aktuelle Normungsvorhaben. In: *Information – Wissenschaft und Praxis*. Band 63, Nr. 1, S. 45-55
- Lauser, Boris; Johannsen, Gudrun; Caracciolo, Caterina, Keizer, Johannes; van Hage, Willem Robert; Mayr, Philipp (2008): Comparing human and automatic thesaurus mapping approaches in the agricultural domain. In: 10th International Conference on Dublin Core and Metadata Applications. <http://arxiv.org/abs/0808.2246>
- Mayr, Philipp & Petras, Vivien (2008): Building a terminology network for search: the KoMoHe project <http://arxiv.org/abs/0808.0518>
- Mayr, Philipp (2010): Information Retrieval Mehrwertdienste für Digitale Bibliotheken: Crosskonkordanzen und Bradfordizing. GESIS-Schriftenreihe 5. Bonn: GESIS. ISBN 978-3-86819-012-0
- Pfeffer, Magnus (2013): Automatic creation of mappings between classification systems. In: *Workshop Klassifikation und Sacherschließung (LIS'2013)*, University of Luxembourg (Vortrag). urn:nbn:de:swb:90-357776
- Reiner, Ulrike (2010): Automatische DDC-Klassifizierung bibliografischer Titeldatensätze der Deutschen Nationalbibliografie. In: *Dialog mit Bibliotheken* 2010/1, 22. Jahrgang, S. 23-29.URN <urn:nbn:de:101-2011012860>

- Reiner, Ulrike (2009): Bewertung von automatisch DDC-klassifizierten Titeldatensätzen der Deutschen Nationalbibliothek (DNB). VZG-Colibri-Bericht 1/2008, VZG. <http://taipan.dyndns.org/~ul/colibri05.pdf>
- Reiner, Ulrike (2008): DDC-based Search in the Data of the German National Bibliography. In: *Gedenkschrift New Perspectives on Subject Indexing and Classification. Essays in Honour of Magda Heiner-Freiling*. DNB, S. 121-129. ISBN 978-3-933641-99-1
- Reiner, Ulrike (2003): VZG-Projekt Colibri Überblick, Stand, Ergebnisse. VZG-Colibri-Bericht 1/2003
- OCLC (2014): 2014 Annual Report to VIAF Council. August 2014 <http://oclc.org/content/dam/oclc/viaf/OCLC-2014-VIAF-Annual-Report-to-VIAF-Council.pdf>
- Ontology Alignment Evaluation Initiative (OAEI). Library Track (2012-2014). <http://oaei.ontologymatching.org/2012/library/>, <http://oaei.ontologymatching.org/2013/library/>, <http://oaei.ontologymatching.org/2014/library/>.
- Soergel, Dagobert (2011): Conceptual foundations for semantic mapping and semantic search. In: F. Boteram, W. Gödert & J. Hubrich (Hrsg.): *Concepts in context*, S. 13-35. Würzburg: Ergon.
- Vrandečić, Denny & Krötzsch, Markus (2014): Wikidata: A Free Collaborative Knowledgebase. In: *Communications of the ACM*, Vol. 57 No. 10, S. 78-85 DOI 10.1145/2629489
- Waidmann, Simone (2014): Erschließung historischer Bestände mittels Crowdsourcing: Eine Analyse ausgewählter aktueller Projekte. In: *Perspektive Bibliothek*. Bd. 3, Nr. 1, S. 59-84. <http://dx.doi.org/10.11588/pb.2014.1.14020>

II. Anlangen

A. Vorhandene DDC- und RVK-Konkordanzen

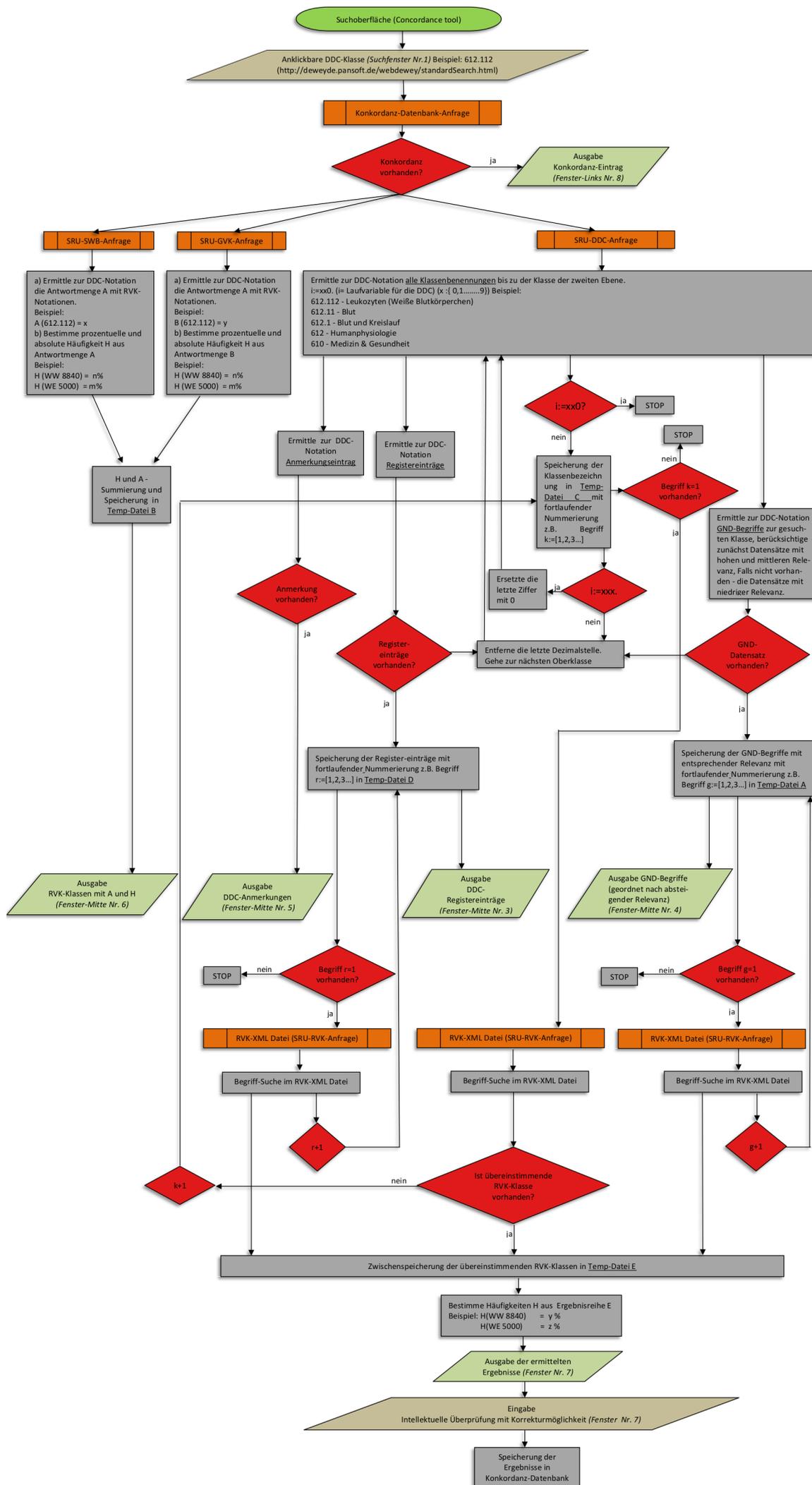
Die folgende Tabelle fasst das Ergebnis einer 2011 durchgeführten Umfrage und weitere Vorarbeiten zusammen (siehe Balakrishnan 2013).

Konkordanz	Fachbereich	Ansprechpartner	Ergebnis
DDC – BK	Chemie	Thomas Hapke TUB Hamburg-Harburg	k.A.
	Politikwissenschaft	Maria Kesting SUB Hamburg	k. A.
	Die Tausend DDC-Klassen der dritten Ebene	Projekt Colibri/DDC Teilprojekt coli-conc U. Balakrishnan Verbundzentrale des GBV	vollständig
DDC – EZB	EZB-Fachgruppen	"	vollständig
DDC – RVK	Die Tausend DDC-Klassen der dritten Ebene	"	vollständig
	Medizin und Gesundheit	"	vollständig
	Recht	"	vollständig
	Philosophie	"	in Arbeit
	Bibliotheks- und Informations- wissenschaften	Prof. Heidrun Wiesenmüller, Fabian Heel HdM Stuttgart	Teilkonkordanz
	Geistes- und Sozialwissenschaften	UB Greifswald	vollständig
RVK – BK	Germanistik, Recht, Pädagogik, Theologie, Politik	Frau Plößnig Universitäts- und Landesbibliothek Tirol	vollständig
RVK – DDC	Biologie, Chemie, Geologie, Mathematik, Paläontologie, Physik	KoMoHe Projekt Dr. Philip Mayr Leibniz Institute for Social Sciences GESIS	Teilkonkordanz
	Psychologie	Manuela Queitsch SLUB Dresden	vollständig

RVK – MSC	Mathematik	CARMEN Projekt Dr. Albert Schröder, Frau Dr. Ines Häusler UB Regensburg	Teilkonkordanz
RVK – PACS	Physik	CARMEN Projekt Dr. Albert Schröder, Frau Dr. Ines Häusler UB Regensburg	Teilkonkordanz
GND – DDC		Criss-Cross Projekt Helga Karg DNB, Frankfurt	vollständig
	Bibliotheks- und Informationswissenschaften	Prof. Heidrun Wiesenmüller; Fabian Heel HdM Stuttgart	Teilkonkordanz
SWD-RVK	Bibliotheks- und Informationswissenschaften	Prof. Heidrun Wiesenmüller; Fabian Heel HdM Stuttgart	Teilkonkordanz
RVK-BK- MSC-PACS	Mathematik, Physik	Veronika Plößnig Universitäts und Landesbibliothek Tirol	Kreuzkonkordanz
DDC – MSC – BK	Mathematik	Bernard Tempel TIB/UB Hannover	Kreuzkonkordanz
RVK – GND	RVK Register	Frau Dr. Ines Häusler UB Regensburg	Teilkonkordanz

B. Algorithmus zur Erstellung von DDC-RVK Konkordanzen

Das im folgenden Ablaufdiagramm dargestellte Verfahren wurde im Projekt *coli-conc* für die automatische Erstellung von Vorschlägen für DDC-RVK-Konkordanzen entwickelt. Das bislang intellektuell durchgeführte Verfahren soll im Projekt als Webservice umgesetzt und im Konkordanz-Tool angeboten werden (siehe AP 4).



C. Vorhandener Prototyp der Benutzeroberfläche

Der Screenshot zeigt die Benutzeroberfläche eines funktionsfähigen Prototypen des Konkordanz-Tools. Derzeit ist nur eine Anbindung an die GND (über lobid.org) und an die RVK (über API der Universitätsbibliothek Regensburg) umgesetzt. Die Komponenten der Vorschlagsfunktion und zur Analyse vorhandener Konkordanzen (im Screenshot in der Mitte) sind erst angedacht.

Cocoda Mapping Tool
Log out

Source Scheme: GND

OR

4036762-9 Lymphozyt

Lymphozyten, Lymphocyt, Lymphzelle, Lymphocyten

◀ **4074195-3** Leukozyt

Use → Look up mappings

4036762-9 → **WF 9880**
WW 9120
WF 9895

Created: 2014-01-01 Clear all Save

Target Scheme: RVK

Search Tools

WW 9120 Lymphsystem (Lymphozyt), Milz, Thymus / Allgemeines

Lymphozyt, Lymphsystem / Vorklinik, Milz / Vorklinik, Thymus / Lymphsystem

◀ **WW 8640** Körperflüssigkeiten und deren Systeme einschließlich deren Bildungsorgane

Add Replace all

VZG-Konkordanz

GND	RVK	Creator	Date
4036762-9	WF 9880 WW 9120	VZG	2014-01-01
4036762-9	WF 9895	VZG	2013-01-01

Catalog Occurrences

Used notation: **4036762-9**

Used database: **GVK**

Results (total) for **4036762-9**: 42

Corresponding notations in **RVK**:

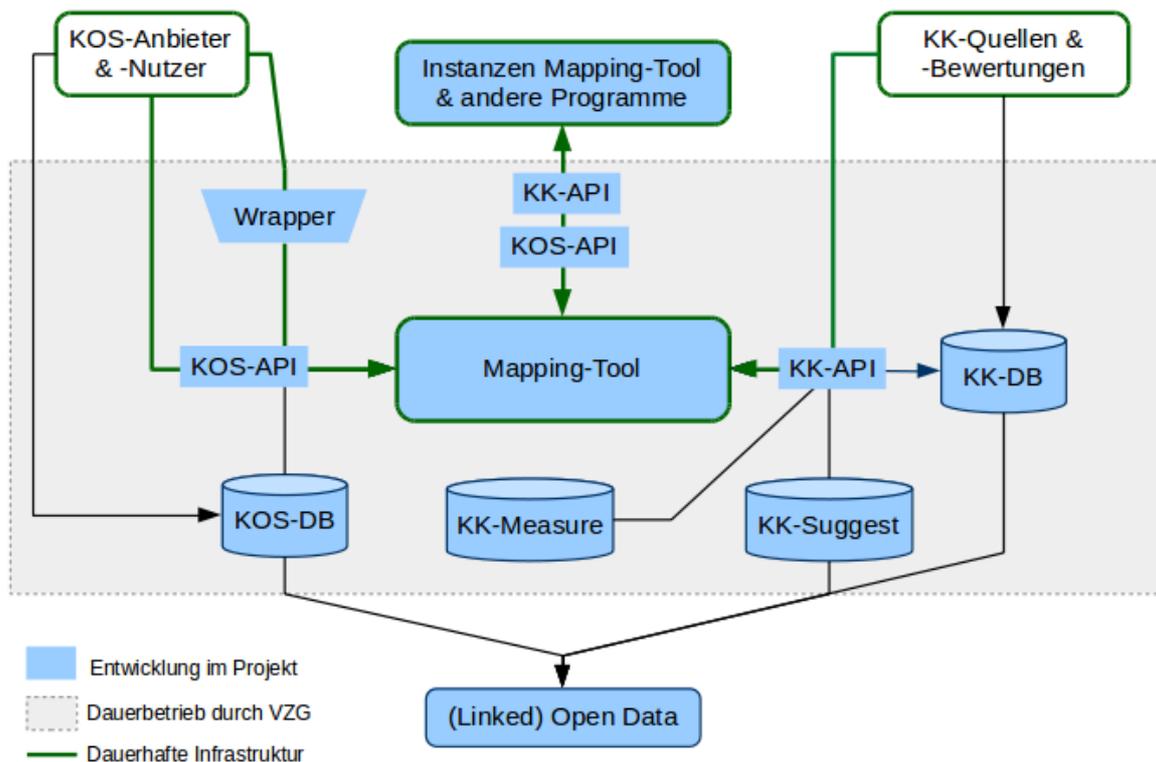
Notation	Hits	% of total
WW 8840	22	52.4 %
YC 2500 - YC 2599	11	26.2 %
WF 9895	8	19.0 %
XG 6700 - XG 6728	1	2.4 %

Top Concepts

- 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

D. Systemübersicht der geplanten Infrastruktur

Das folgende Diagramm gibt eine Übersicht der geplanten technischen Infrastruktur. Wichtigster Beitrag zum Aufbau einer dauerhaften Infrastruktur ist die Schaffung einheitlicher Schnittstellen bzw. Wrappern für den Zugriff auf Wissensorganisationssysteme (KOS-API) und Konkordanzen (KK-API), so dass sowohl verschiedene Instanzen des Konkordanz-Tools als auch vorhandene KOS-Software (iQVoc, Poolparty...) und -Dienste (lobid.org, ZBW-API, RVK-API, Wikidata...) sowie Quellen für Konkordanzen und Bewertungen gegenseitig auf Inhalte zugreifen können.



H. Letters of Intent

Es liegen Bestätigungen zur geplanten Zusammenarbeit mit folgenden Institutionen vor:

- Leibniz Institut für Sozialwissenschaften (GESIS), Köln
- Wikimedia Deutschland – Gesellschaft zur Förderung Freien Wissens e.V., Berlin
- Deutsche Zentralbibliothek für Wirtschaftswissenschaften (ZBW)
- Leibniz Informationszentrum Wirtschaft, Kiel
- Universitätsbibliothek der Freien Universität Berlin
- Universitätsbibliothek St. Gallen, Schweiz
- Universitätsbibliothek Regensburg
- Universitäts- und Landesbibliothek Tirol, Innsbruck