Mapping Project coli-conc

Progress, learning & next steps

U. Balakrishnan, J. Agne, J. Voß

Content

- Aim & Project Start

- Partners

- coli-conc-Key Objectives & Approaches

- Work Packages

- Survey

- JSKOS Dataformat

- Literature

Aim & DFG Project start

- Aim: Infrastructure to aid semi-automatic creation and management of mappings between library KOS and enable a facile access to the same.

- Project start: 15. December 2015

Partners

VZG

www.gbv.de

coli-conc-Key Objectives & Approaches

Effective Creation and Management of Mappings
- Development of a Mapping tool
- Provision of KOS and their Mappings
- Creation or adoption of concordance algorithms
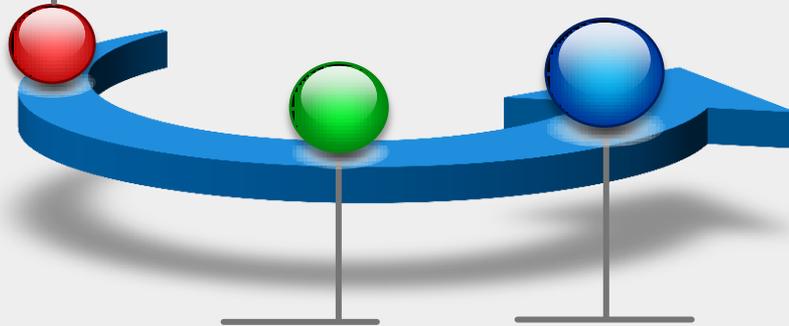
Improvement of the Quality of the Mappings
- Development and Implementation of measures for Quality Assessment
- Involvement of Experts and User groups

Facilitate Use and Exchange of KOS and their Mappings
- Collection of KOS and their existing Mappings
- Provision of KOS and their Mappings
- Documentation and Outreach

Work Packages

Evaluation and Study

- Current status of KOS and their mappings
- Existing software for KOS management & creation of mappings

Development of a Web Interface

- Representation & navigation of KOS
- Compilation and processing of mappings
- Display and workspace on auto-generated mapping candidates
- Usability-tests of the web interface

Concordance Algorithm

Provision

- Storage and provision of KOS (KOS-DB)
- Management of concordances (KK-DB)
- Integration with the existing KOS-software, gateways and cataloging tools
- Testing and documentation of know-how for installation of the application in cloud services
- Publishing mappings as linked open data

Quality Assessment

- Evaluation of existing standards and know-how for quality assessment
- Conception and testing of crowd-sourcing strategy
- Development of a server component to assess and provide quality measurements

Outreach

- Documentation on KOS and mapping tool
- Hold Workshops with experts and users
- Conception and implementation of authentication and authorization techniques
- Conception of crowd sourcing approach
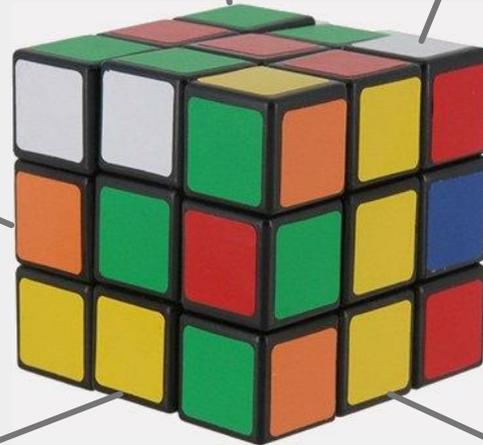
survey

General information

- Aim
  - Evaluation of …

    … the distribution of KOS

    … the software of KOS

    … the concordance projects

- Distribution
  - INTEBIB, KIM, ÖBIB,

    FAGEI, EGSE, RVK-Mailingliste
  - Facebook, Twitter
  - hardcopy (library congress Leipzig 2016)

- Tool: qset

- Timeline: 08.02.2016 – 08.04.2016

Participants

over 200

Questionnaires

KOS:

- Distribution of KOS
- Software for KOS management
- Reasons for choosing KOS
- Purposes of KOS
- How subject indexing is done
- KOS system changes

Mapping Projects:

- Mapping projects
- Building concordances
- KOS
- Subjects
- Types of concordances
- Mapping relationships
- Mapping types
- Mapping tool

Results Part I - KOS

KOS in Germany I – Survey 2016

2016

classification systems

subject headings & thesauri

www.gbv.de
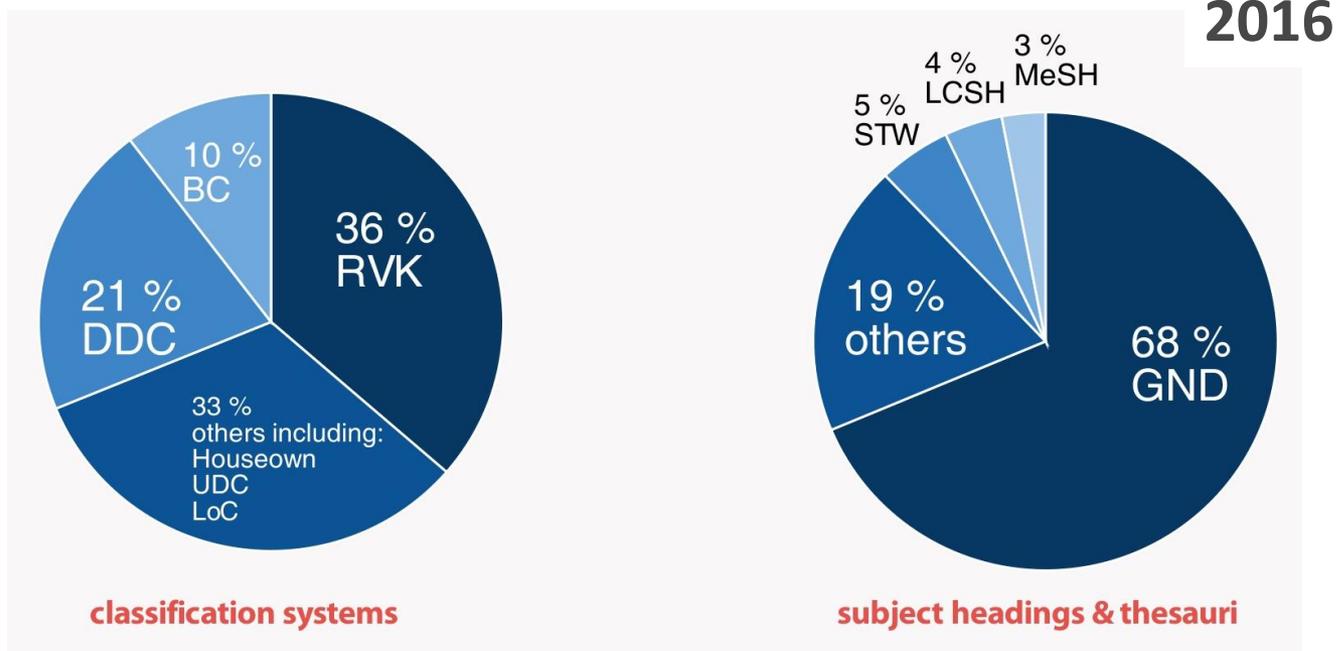
KOS in Germany II – Updated List of KOS

VZG Terminologie Webapplication

Universal Classification systems	No. of classes
UDC (Universal Decimal Classification)	65.000 classes
DDC (Dewey Decimal Classification)	10 main classes and 44.000 classes in total
RVK (Regensburg Classification)	34 main classes and 855.000 classes in total
BC (Basic Classification)	89 main classes and 2100 classes in total
LCC (Library of Congress Classification)	21 main classes
GHB (Overall University Library Classsification)	36 main classes
ASB (General Classification of the Public Libraries)	23 main classes 2200 subclasses
ÖSÖB (Austrian Public Library Classification)	
SfB (Classification for (public) libraries)	30 main classes and 14400 subclasses
KAB (Classification of General Libraries)	2700 classes
Subject classification	
DDC-Sachgruppen der DNB	10 main classes with 94 subclasses
MSC (Mathematics Subject Classification)	87 main classes
PACS (Physics and Astronomy Classification Scheme)	10 main classes
FKDigBib (Subject Classification for Digital Libraries)	10 main classes
KfM (Classification of Music libraries)	ca. 800 classes
Classification of the German Patent and Trademark Office	
SEB (Classification of Protestant Libraries)	
Non-fictional Classification of Catholic Public Libraries	
KfKJ (Classification of Children and Youth Libraries)	< 1000 classes
SSD (Classification of the Public Library Duisburg)	
ESSB (Unity Classification of Libraries of South Tyrol)	16 main classes
Systematic of the library of the region Stade	
Classification for Social Studies - GESIS	
Classification of the finance court Cologne	
FIV subject classification	
FIV classification of the region	
Psychoanalytic classification	
Journal of Economic Literature Classification System	
OSKAR (Online Systematic for maps)	
Classification of the institute of contemporary history	Ca. 54 main classes
SMM Systematik Musikbibliotheken	
FDC (GFDC Global forest decimal classification)	
IxTheo (Index theologicus)	

Subject Classifications at the Universities	
TUM-classification (Science and technology classification of the TU Munic)	52 classes each with 999 notations
Subject classification of the University library Duesseldorf	45 classes
Bremer classification of the State and University library Bremen	ca. 57 main classes
GOK (Goettingen Online Classification)	ca. 33 main classes
Systematik des Göttinger Band-Realkatalogs	
Subject classification University library Trier	36 main classes
Technical University Dortmund	28 main classes
University library Paderborn	26 main classes
University library Marburg	35 main classes
University library Bonn	24 main classes
University library Heidelberg	22 main classes
Subject classification and nomenclature of individual languages Library of the Institute of General Linguistics at the Uni Münster	23 main classes
Thesauri	
Standard-Thesaurus Wirtschaft von der ZWB	6.000 Terms and notations
APA (thesaurus of the American Psychological Association)	
MeSH (Medical Subject Headings)	
CAB thesaurus (thesaurus of the Centre for Agriculture and Biosciences)	
LCSH (Library of Congress Subject Headings)	
Rameau (Répertoire d'autorité-matière encyclopédique et alphabétique unifié = Encyclopaedic directory and unified alphabetical subject authority)	
RSWK (Rules for the keyword index)	
ACM (Computing Classification System)	

http://coli-conc.gbv.de/terminologies/
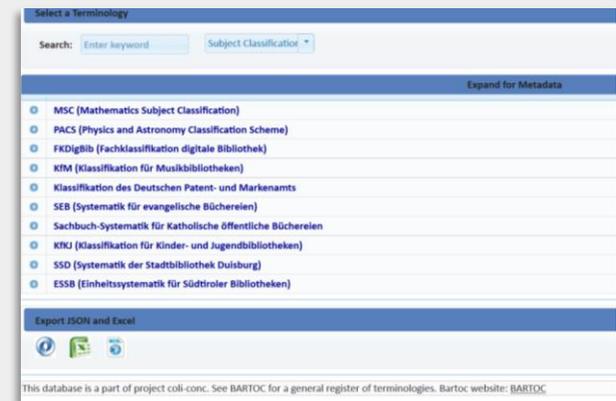
KOS in Germany III – Surveys 1999 & 2011

1999

- 33 % houseown
- 33% RVK
- 15% BC
- 8% GHBS NRW
- UDC, DDC 5%
- Others 6%

2011

- RVK 46%
- DDC 26%
- Others 17%
- BC 8%
- UDC 3%
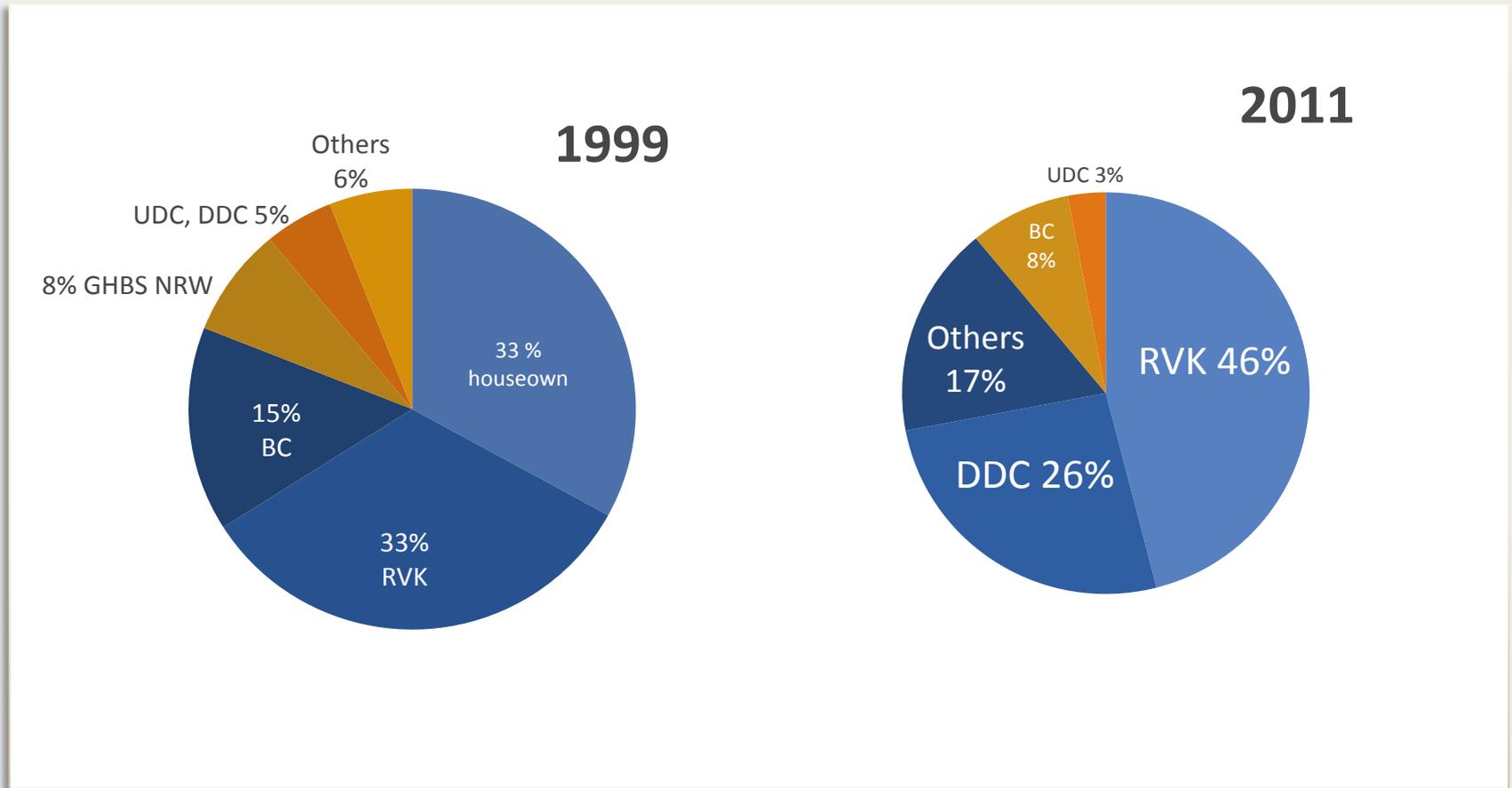
A look at Dewey
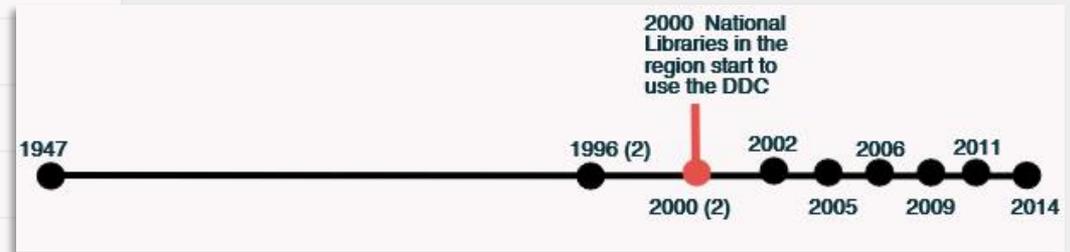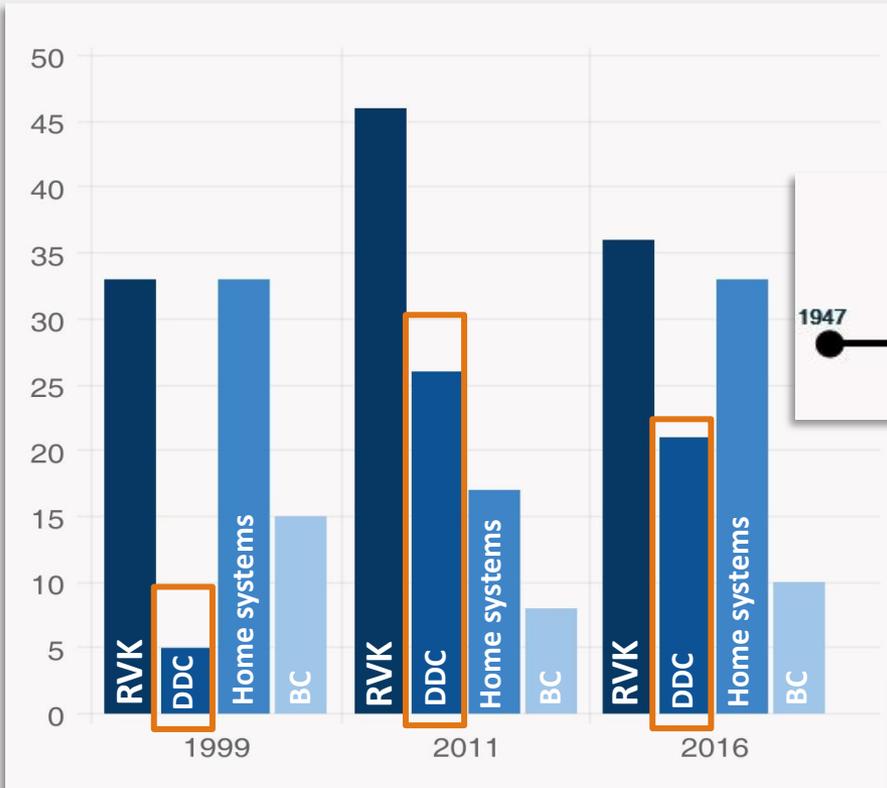
KOS System Change

system change?

8 % YES
11% maybe
81 % NO

to

10 % undecided
5 % DDC
5 % Thesaurus
5 % BC
15 % others
60 % RVK
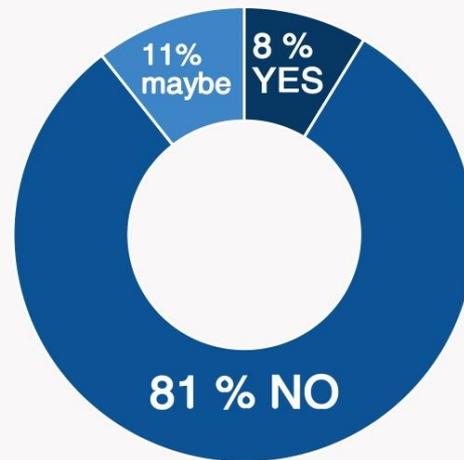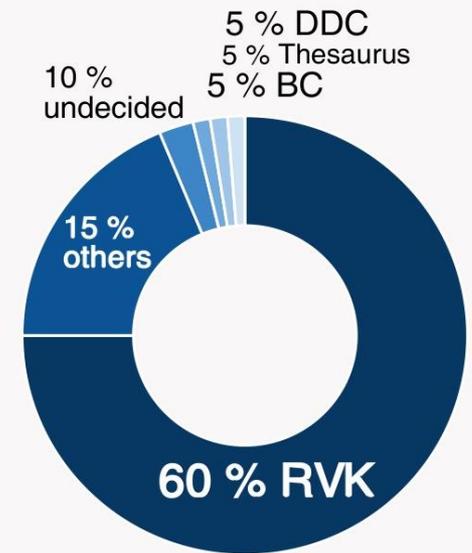
Software used for KOS Management

PICA

Aleph

other: OSIRIS/BREWIS, DSWORM, ABACUS, E-LIB, Alephino

Name:
Pica
Aleph
aDIS/BMS
URICA
ALMA
SydneyPLUS/Lucidea
Allegro
Eigenentwicklung
DSWORM
OSIRIS/BREWIS
ABACUS
E-LIB
Alephino
DSpace
BibliothecaPlus
Office
Geocom Bibliothek (Lotus Notes basiert)
FileMaker (Eigenentwicklung)
SISIS-SunRise (lokal)
IFIS
HohSearch
STAR

Subject indexing - How is it done?


96 %
intellectually
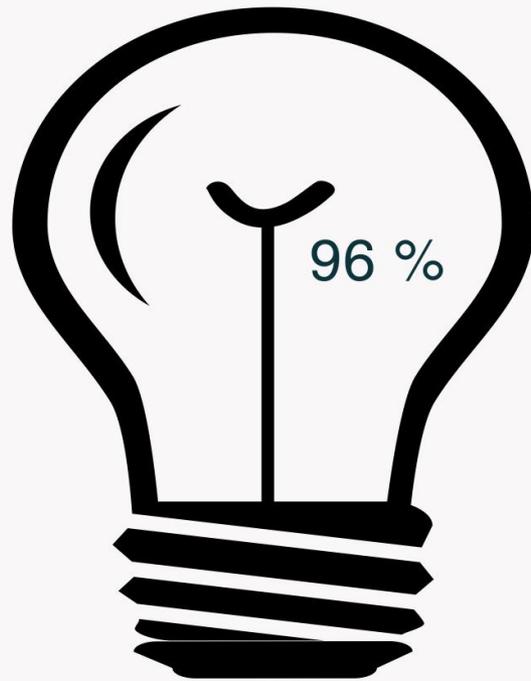
4 %
automatically

Software:

- SisisSunrise library system
- SIKOM-interface
- InteligentCapture
- Averbis Extraction Platform
- Semiautomatic Digital Assistent by Eurospider

Results Part II - Mapping Projects

Mappings I – Projects & Concordances
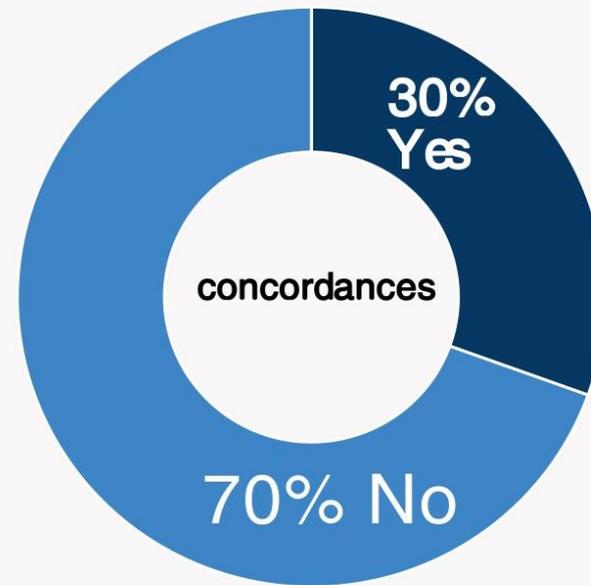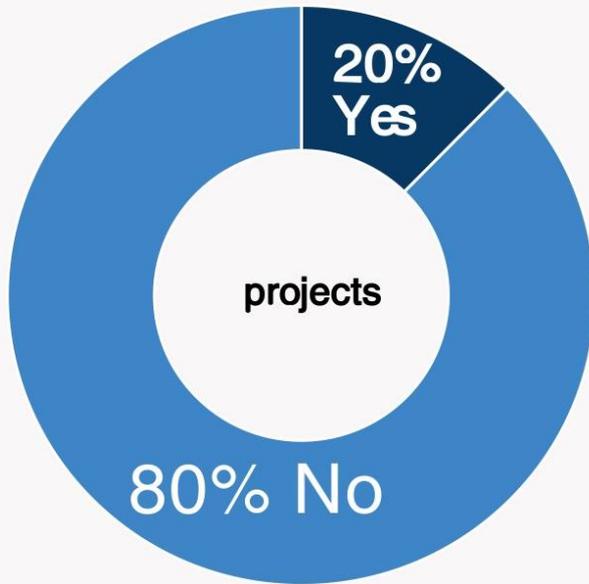
Mappings – Processes & Types

Classification systems

mostly: manual/intellectual, unidirectional mappings

Thesauri

few: semi-/automatic mappings

 (tools only for thesauri: OSIRIS, Amalgame, Decision Trees)

Mappings – KOS mapped to Dewey
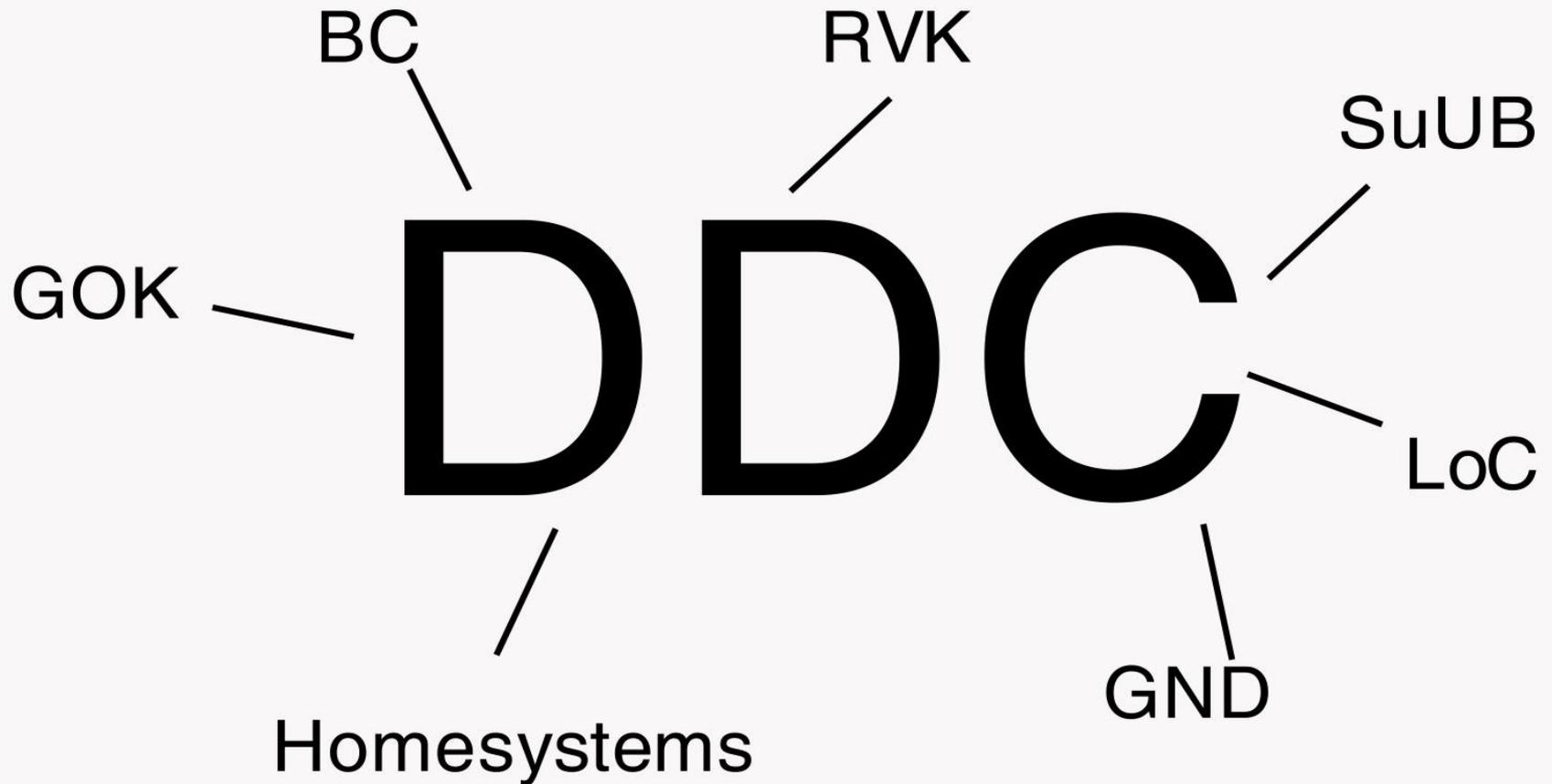
Concordances

KOS and topic	creator	No. of Mappings
DDC to ...		
DDC-BC Chemistry	unknown	100
DDC-BC politics	SUB Hamburg	851
DDC-RVK (DDC Class 100)	VZG	754
DDC-RVK (DDC first 1000 classes)	VZG	2065
DDC-RVK Library and Information science	HdM Stuttgart	385
DDC-RVK Ethics	VZG	224
DDC-RVK Medicine	VZG	2814
DDC-RVK Law	VZG	2268
DDC-BK (DDC first 1000 classes)	VZG	1376
RVK to ...		
RVK-BC Law	University and Regional Library of Tyrol	16836
RVK-BC Economics	University and Regional Library of Tyrol	1042
RVK-DDC Ethics	VZG	27
RVK-DDC Philosophy and Psychology	VZG	358
RVK-GND	UB Regensburg	94196
RVK-BK German studies	University and Regional Library of Tyrol	13249
SWD to ...		
SWD-DDC Library and Information science	HdM Stuttgart	462
SWD-RVK Library and Information science	HdM Stuttgart	396

currently total: 137403

coli-conc Mapping Database:

http://coli-conc.gbv.de/concordances/

Conclusion

1. The use of the DDC has certainly increased since 2000, however not significantly

2. The RVK is consistently well represented

3. Still there is a huge number of libraries using home classification systems that do not strive for a change or a uniform standard

4. Several software are available, for management of KOS, however methods, standards and tools for creating mappings between KOS, in particular between classification systems lacking

JSKOS data format

Objectives

1. Creation and Manage of Concordances and KOS
2. Improvement the Quality of the Mappings
3. **Facilitate Use and Exchange of KOS and their Mappings**

This requires

- I an easy-to-use access method (JSKOS-API)
- I an easy-to-use data format (JSKOS)

The JSKOS data format

- Based on JSON(-LD) $\Rightarrow$ compatible with SKOS/RDF
  but easier to use, especially in web applications

- Adds expression of concordances, mappings, and registries
  (which were lacking in pure SKOS or other formats)

- Reference implementations in PHP, JavaScript & Java

The JSKOS data format

- Based on JSON(-LD) ⇒ compatible with SKOS/RDF
  but easier to use, especially in web applications

- Adds expression of concordances, mappings, and registries
  (which were lacking in pure SKOS or other formats)

- Reference implementations in PHP, JavaScript & Java

```json
{
  "uri": "http://dewey.info/class/641.5/e23/",
  "notation": [ "641.5" ],
  "prefLabel": { "en": "Cooking", "de": "Kochen" },
  "narrower": [ ... ],
  ...
}
```

JSKOS-API

Query KOS and mappings to include in other applications

- cocoda mapping tool
- subject indexing
- . . .

REST-API

- ?notation=614.5&language=en,de
- ?search=Cook&format=suggest
- . . .

Current state

Done enough to build on

- JSKOS specification http://gbv.github.io/jskos/
- Several Open Source implementations (see homepage)
- GND and RVK expressed in JSKOS

In progress

- JSKOS-API specification
- Implement JSKOS-API in existing KOS software
- Express DDC in JSKOS, lacking dewey.info :-(

Thank you!