Precombination vs. Precoordination

Comparing LSCH and RSWK
Two subject heading languages

Library of Congress Subject Headings (LCSH):

• prevalent Anglo-American indexing language
devolved by the Library of Congress since 1898
• inspired many other subject heading languages
e.g. the French RAMEAU

Regeln für den Schlagwortkatalog (RSWK):
(Rules for subject catalogs)

• indexing language of the German speaking countries
  used in Germany, Austria and Switzerland
• first edition 1986
  current: 3rd ed. 1998 (last revised in 2010)
  major revision impending
Agenda

1. Some basic principles
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
Agenda

1. **Some basic principles**
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
Precombination vs. precoordination

LCSH:
Academic libraries—Collection development

RSWK (with English translation):
Wissenschaftliche Bibliothek ; Bestandsaufbau
Academic library ; Collection development

looks superficially similar
but: underlying principles completely different
LCSH:
Academic libraries—Collection development

Two elements which have been precombined in advance („glued together“) to form one single heading

RSWK:

Wissenschaftliche Bibliothek ; Bestandsaufbau
Academic library ; Collection development

Two separate elements which are precoordinated („put together“) during the process of indexing according to the topic of the resource in hand
Authority records

• LCSH: one single authority record for a complex concept, comprising two aspects

150 ___ |a Academic libraries |x Collection development
550 ___ |w g |a Collection development (Libraries)

one large building block
• **RSWK**: two authority records
  each for a simple concept: small building blocks

**Academic library**

150 Wissenschaftliche Bibliothek
450 Wissenschaftliche Bibliotheken
550 !208865586!Bibliothek$4obal

**Collection development**

150 Bestandsaufbau
450 Erwerbungspolitik$VQuasisynonym
550 !211955949!Erwerbungsrichtlinie$4vbal
550 !208911774!Erwerbung$4vbal
Structure of authority files

- many authority records needed in LCSH
  each complex concept needs its own authority record

**LCSH:** nine authority records for the following concepts

- Academic libraries—Collection development
- Academic libraries—Interlibrary loans
- Academic libraries—Reference services
- Public libraries—Collection development
- Public libraries—Interlibrary loans
- Public libraries—Reference services
- Medical libraries—Collection development
- Medical libraries—Interlibrary loans
- Medical libraries—Reference services
RSWK: only six records needed to express the concepts:

- Wissenschaftliche Bibliothek (Academic library)
- Öffentliche Bibliothek (Public library)
- Medizinische Bibliothek (Medical library)
- Bestandsaufbau (Collection development)
- Leihverkehr (Interlibrary loans)
- Auskunftsdienst (Reference services)

- single concept headings can be freely combined to form the necessary subject heading strings, e.g.:

  - Wissenschaftliche Bibliothek ; Bestandsaufbau
  - Öffentliche Bibliothek ; Bestandsaufbau
  - Medizinische Bibliothek ; Bestandsaufbau
  - etc.
Rules for combination in RSWK

Fixed citation order according to primary categories:
1. persons (p)
2. geographic aspects (g)
3. topical aspects (s)
4. temporal aspects (z)
5. form aspects (f)

- Note: indicators for categories usually not shown in library catalogs
Precoordination in LCSH

- precombination supplemented by precoordination
  e.g. geographic subdivisions and „free-floating“ subdivisions, which can be freely added to precombined headings (note: no full authority record in these cases)

Examples:

Academic libraries—Collection development—United States—History—20th century

Public libraries—Reference services—Handbooks, manuals, etc.
Introduction of new topics

RSWK:
- in most cases no need for new authority records
  typically the necessary single concept headings are already there and only have to be combined in a new way

LCSH:
- often no suitable authority record exists
  or can be built by using free-floating subdivisions etc.
- new headings difficult and time-consuming to create

  typical solution: combination of several existing headings although each of them is broader than the topic of the resource in hand
Building Area Studies Collections

Edited by Dan Hazen and James Henry Spohrer

Topic: Development of collections for area studies (Africa, Southeast Asia, Latin America etc.) in libraries
• **LCSH:** *three headings, each of them fairly broad*

Bibliothek; Kulturraumforschung; Bestandsaufbau; Aufsatzsammlung

Library; Area studies; Collection development; Essays

• **RSWK:** *one subject headings string which matches the topic exactly*
Agenda

1. Some basic principles
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
User understanding

- LCSH/RSWK: sometimes rather longish constructs can easily consist of more than three bits of information.
- RSWK: headings are simply put one after the other; no additional means of expressing relationships.
- LCSH: makes use of prepositions and conjunctions close to natural language, more expressive and easier to understand than structured headings.

Examples:

Libraries and children with mental disabilities
Librarians in motion pictures
Cows on postage stamps
Understanding of structured strings?

<table>
<thead>
<tr>
<th>Personal name:</th>
<th>Eberlein, Johann Konrad.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main title:</td>
<td>Harald de Bary : Leben und Werk / von Johann Konrad Eberlein ; mit einem Beitrag von Ilonka Czerny und dem vollständigen Werkverzeichnis = Harald de Bary : life and work / by Johann Konrad Eberlein ; with a contribution from Ilonka Czerny and a complete catalogue of his work.</td>
</tr>
<tr>
<td>Published/Created:</td>
<td>Frankfurt am Main : Dielmann, c2006.</td>
</tr>
<tr>
<td>Description:</td>
<td>223 p. : ill. (chiefly col.) ; 34 cm.</td>
</tr>
</tbody>
</table>
| ISBN:         | 9783866381230  
|               | 3866381239     |
| Parallel title: | Harald de Bary : life and work |
| Related names | Bary, Harald de, 1935- |
|               | Abstract expressionism --Germany --Catalogs. |

- **Harald de Bary:**
  exponent of a type of abstract art called „Informel“ or „Informal art“ (French: „art informel“)
Bachelor thesis (Sabrina Stutz):

- only the subject headings were shown to students
  test persons were then asked what the book is about

- results for this example:
  - several test persons did not understand that the book
    is about Harald de Bary
  - some test persons thought that the three strings referred
    to three different books
Should we re-think presentation?

• present several topics in a clearer way

<table>
<thead>
<tr>
<th>Topic 1:</th>
<th>(...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic 2:</td>
<td>(...)</td>
</tr>
<tr>
<td>Topic 3:</td>
<td>(...)</td>
</tr>
</tbody>
</table>

• break up strings in several facets, e.g.

<table>
<thead>
<tr>
<th>Topic 2:</th>
<th>Person treated: Bary, Harald de</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Form of treatment: Biography</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Topic 3:</th>
<th>Person treated: Bary, Harald de</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Form of treatment: Catalogue raisonné</td>
</tr>
<tr>
<td></td>
<td>Period covered: 1955-2005</td>
</tr>
</tbody>
</table>
Agenda

1. Some basic principles
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
Strengths and weaknesses

- **LCSH:** *strong on browsing, weak on keyword searching*
- **RSWK:** *weak on browsing, strong on keyword searching*

Browse index

Number of entries

- **LCSH:** *headings often fairly general*
  - *reasonable number of different headings in the index, often several titles with the same heading*
- **RSWK:** *very specific strings*
  - *very many different strings in the index, often only one title for each string*
<table>
<thead>
<tr>
<th>#</th>
<th>Hits</th>
<th>Headings (Select to View Titles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>Stuttgart (Germany)--Biography.</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Stuttgart (Germany)--Biography--Portraits.</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>Stuttgart (Germany) Bosch-Areal</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.--Catalogs.</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.--Designs and plans.</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.--Designs and plans--Exhibitions.</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>Stuttgart (Germany)--Buildings, structures, etc--Exhibitions.</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.--Guidebooks.</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>Stuttgart (Germany)--Buildings, structures, etc.--Pictorial works.</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>Stuttgart (Germany)--Census, 1987--Maps.</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>Stuttgart (Germany)--Church history.</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>Stuttgart (Germany)--Church history--Exhibitions.</td>
</tr>
</tbody>
</table>

*extract from LC’s browse index*
extract from the browse index of the Southwest German library network (SWB)
Additional entry points:

- LCSH: covered by structural references
  only possible if there is an authority record

But:

- Academic libraries—Austria
  no entry point under „Austria“

- Austria—Economic conditions
  no entry point under „Economic conditions“
• RSWK: covered by permutations
  order of the headings in a string is changed in order to bring each significant heading to front position

Stuttgart ; Architektur ; Geschichte 1875-1924
*Stuttgart ; Architecture ; History 1875-1924*

second, permutated string:

Architektur ; Stuttgart ; Geschichte 1875-1924
*Architecture ; Stuttgart ; History 1875-1924*

But:

• no longer obligatory since 2010
  was also never done consistently in former times

• alternatives need to be implemented
  e.g. KWOC index
Keyword searching

Different data models:

• **German-speaking countries**
  *title records are linked with authority records, both headings and see references can be used in keyword searching*

• **Anglo-American world**
  *mostly no links from title records to authority records: only headings can be searched, but not see references*

  *general technical problem, which will hopefully be overcome by technical means in the near future*
Structural problems in LCSH:

- "see references" for synonyms
  stored in authority records for basic concepts only

```
150 __ |a Social sciences
450 __ |a Behavioral sciences
450 __ |a Human sciences
450 __ |a Sciences, Social
450 __ |a Social science
450 __ |a Social studies
550 __ |w g |a Civilization

150 __ |a Digital libraries
450 __ |a Digital curation
450 __ |a Digital media collections
450 __ |a Digital media libraries
450 __ |a Electronic libraries
450 __ |a Electronic publication collections
450 __ |a Electronic publication libraries
450 __ |a Electronic text collections
450 __ |a Virtual libraries
550 __ |w g |a Libraries
550 __ |a Information storage and retrieval systems
550 __ |a Web archives
```
• precombined headings
  no „see references“ for synonyms

150 __ |a Social sciences |x Digital libraries
550 __ |w g |a Digital libraries

• geographic or free-floating subdivisions
  no authority records, i.e. no references possible

these problems are unknown in RSWK
due to its different structure
Agenda

1. Some basic principles
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
Seven facets aimed specifically at the browsing of poems

http://www.poetryfoundation.org
Thanks to Debora Shon for this great example!
Some basic points about facets:

- **specific vs. universal facets**
  
  "poetic terms" or "occasion" specific to a certain area, but there are also universal facets like place and time

- **number and presentation of values**
  
  Facets make most sense if the number of different values is not too large (e.g. "occasion": only 11 values) and the values are well-arranged (e.g. hierarchically as in "poetic terms")

building of facets from RSWK and LCSH should concentrate on universal dimensions of time, place and form
Faceting LCSH/RSWK

RSWK:

• has "built-in" facets
  e.g. person headings, geographic headings, form headings, time headings

• but: usually only one facet for subject headings
  e.g. University Library of Augsburg: all kinds of headings presented in the same drill-down facet

https://opac.bibliothek.uni-augsburg.de/InfoGuideClient.ubasis/start.do?Login=iguba
LCSH:

- complex headings must first be split up *in order to create facets*

- FAST project (OCLC)
  
  *Faceted Application of Subject Terminology*

  United States—Civilization—Italian influences—History—20th century—Sources

*reworked in FAST as:*

  Geographic: United States
  Topical: Civilization—Italian influences—History
  Period: 1900-1999
  Form: Sources

- there are also different attempts at creating facets
  
  *e.g. Endeca catalog of NCSU Libraries*
Time facet (Endeca):

- more normalization needed using FAST headings would help
- too many different values if presented in a facet at all, it would be better to have broader, yet more regular units (e.g. only centuries or decades)
- only explicit years are used there are also cases like e.g. „Art, Early Christian“ or „Punic wars“, where the time information is hidden/implicit

http://www.lib.ncsu.edu/catalog/
Time headings in RSWK:

- even more manifold, as exact years are given, e.g.
  Geschichte 1904-1912
  Geschichte 1892-1929
  Geschichte 1907

  would all be relevant for somebody interested in the time span „1900-1910“

- could be solved by a special algorithm
  which works out the relevant results for every query;
  could be presented as a time bar instead of a facet
  (a concept for this has already been developed)
Region facet (Endeca):

- more normalization needed
e.g. “Boston“ (place as geographic subdivision) vs. “Boston (Mass.)“ (place as main heading), using FAST headings would help

- no hierarchical display
„Europe“, „England“ and „London“ in the same list

- only explicit place information
geographic information about e.g. persons is not covered
Geographic facet based on RSWK

- **two protoypic implementations**
  University Library of Mannheim
  University Library of Heidelberg

- **based on country codes in authority records**
  *hierarchically structured codes:*
  continent – country – (federal state or canton)
  e.g. XA-DE-BW:
  Europe – Germany – Baden-Württemberg

- **country codes are stored in many records**
  *not only in geographic headings, but also in records for persons, corporate bodies, buildings, historic events etc.*

  In retrieval, the recall is much better when using the codes instead of geographic names
Black Forest

French Revolution

Heidrun Wiesenmüller
Stuttgart Media University

043 XA-DE-BW
065 19.1b
083 T2--4346$t2007-01-01
151 Schwarzwald
550 !209582146!Mittelgebirge$4obin

043 XD-US
065 7.14p;16.5p
083 973.932092$4d3
100 Obama, Barack
375 m
400 Obama, Barack Hussein
400 Übämä, Bäräk
500 !302386661!Obama, Michelle*1964-*$4bezf$vEhefrau
500 !178890049!Obama, Auma*1960-*$4bezf$vSchwester
548 1961$4datl
548 04.08.1961$4datx
550 !209069961!Politiker$4berc
550 !208980806!Jurist$4beru
550 !209494697!Präsident$4beru
551 !209559403!Honolulu, Hawaii$4ortg
670 LCAuth
678 $b44. Präsident der USA seit 20.01.2009,
Friedensnobelpreisträger 2009

043 XA-FR
065 16.5
083 940.27$d1$t2009-11-23
083 944.04$d3$t2009-11-23
150 Französische Revolution
450 Revolution$g1789
451 Frankreich$xRevolution$g1789
548 1789$b1799$4datb
551 !208922857!Frankreich$4geoa
667 Beispiel in RSWK 3. Aufl.
670 M
679 1789-1799
• Geographic facet in Mannheim
  short version (left) and full version (right)

http://www.bib.uni-mannheim.de/133.html
Agenda

1. Some basic principles
2. Presentation
3. Browsing and searching
4. Facets
5. Conclusion
Comparing LCSH and RSWK

• radical structural differences between the systems
  very instructive to note and explore them

• problems are partly similar, partly very different
  often it can help to look at the solutions
  of the other subject heading language

• browsing and searching
  RSWK needs to improve on browsing,
  LCSH needs to improve on keyword searching

• presentation and faceting
  should be further developed in both systems
References:

http://www.ub.uni-heidelberg.de/archiv/12555/

Thank you for your attention!

wiesenmueller@hdm-stuttgart.de