Esther Scheven, Deutsche Nationalbibliothek

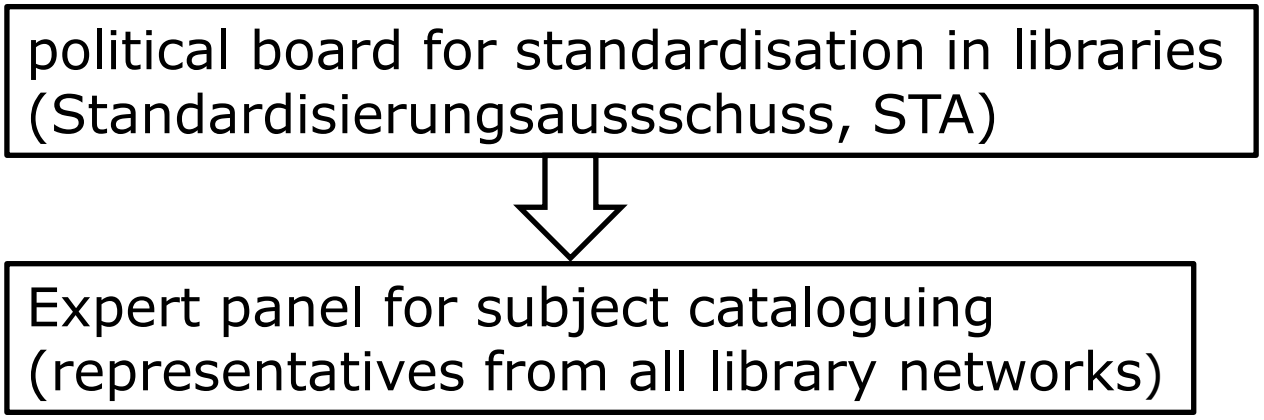
Principles of subject indexing – conclusions from a review by the German expert panel for subject cataloguing

DEUTSCHE NATIONAL BIBLIOTHEK

Subject cataloguing in Germany, Austria and Switzerland („D-A-CH“)

Libraries and library networks do subject indexing cooperatively

Organisation:



Common rules: „Regeln für den Schlagwortkatalog (RSWK)“

Common authority records: Integrated Authority File (GND)
(until 2012: Subject Heading Authority File, SWD)

DEUTSCHE NATIONAL BIBLIOTHEK

Common rules for subject cataloguing: Regeln für den Schlagwortkatalog – **RSWK**

History: Implementation of data processing in cataloguing in Bavaria in the 1980s.

Early partners: Deutsche Nationalbibliothek, Bayerische Staatsbibliothek, university libraries in Bavaria

Editions:

- First edition: 1986
- 3rd edition 1998: loose-leaf-edition with 7 supplements
- Electronical edition 2010 based on the 7th supplement urn:nbn:de:101-2012053100

DEUTSCHE NATIONAL BIBLIOTHEK

Regeln für den Schlagwortkatalog – **RSWK** (Rules for subject cataloguing)

RSWK is based on the German standards for thesauri

Rules for all kinds of entities (persons, corporate bodies, conferences, places, works, topical headings, biological nomenclature, events etc.)

Instructions for cataloguing the form of work
Instructions for cataloguing the temporal aspect

Single subject headings are put together while indexing; this string represents the content of the document

→ A very flexible system (no prefabricated strings as in LCSH)

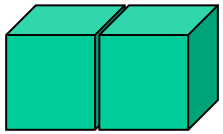
DEUTSCHE NATIONAL BIBLIOTHEK

Regeln für den Schlagwortkatalog – **RSWK** (Rules for subject cataloguing)

Difference between RSWK and LCSH

LCSH:

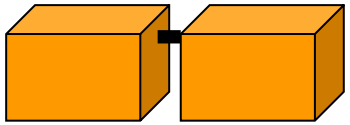
Academic libraries—Collection development



Big, prefabricated building block of several tightly connected elements

RSWK:

Wissenschaftliche Bibliothek ; Bestandsaufbau
Academic libraries ; Collection development



Several small building blocks are put together during indexing according to the content of the resource

DEUTSCHE NATIONAL BIBLIOTHEK

RSWK and the authority file (SWD, GND)

- Simultaneous to the development of RSWK:
  development of a common authority file for subject headings called SWD
- SWD: authority records are created when needed but the principles of a thesaurus are observed
- Since 2012: GND (Gemeinsame Normdatei = Integrated Authority File) instead of SWD

GND (and former SWD) is a hybrid file:
A name authority file for names like persons, corporate bodies, places etc.
A thesaurus for topical headings

DEUTSCHE NATIONAL BIBLIOTHEK

Some information about GND

Before GND:
Different authority files for entities acting as agents used in **descriptive cataloguing**:
persons – name authority file for persons (PND)
corporate bodies – name authority file for corporate bodies (GKD)

Subject authority file for all entities used as subjects in **subject cataloguing** (SWD)

Duplicate records for the same entities in different authority files created by different rules which lead to different access points

DEUTSCHE NATIONAL BIBLIOTHEK

Some information about GND

Aims of GND:

Only one record for the same entity (no parallel authority records for descriptive and subject cataloguing any longer)

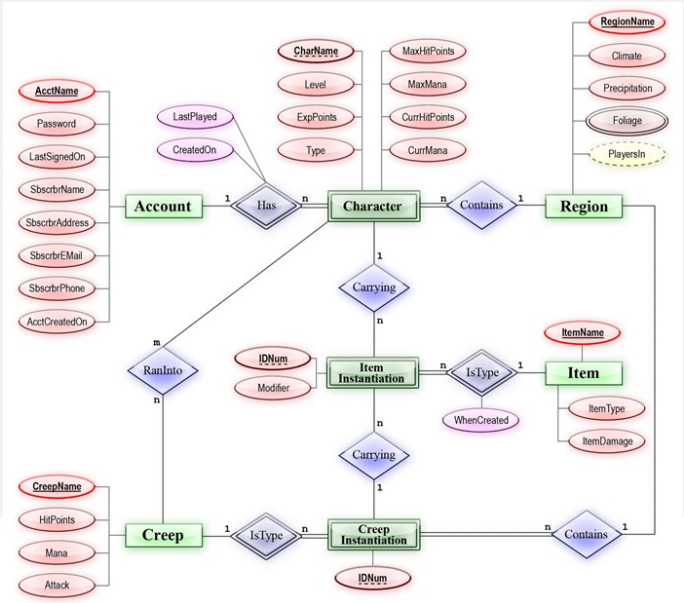
Consequences:

- Same rules for descriptive and subject cataloguing are needed for such entities
  entities affected: entities of group 2 of FRBR
  (persons, corporate bodies including places, families)
  entities of group 1 of FRBR (work, expression etc.)
  common rules since 2014/2015: RDA
- Same authority records for places acting as agent
  needed in descriptive cataloguing and the geographical space
  needed in subject cataloguing (deviation from FRBR)

Modern data model

GND – data model

- Entity–Relationship–Model (ERM)
- Entities have attributes and relationships to other entities
- The types of relationship are specified by codes
- Modular data structure
- GND itself is a semantic web



DEUTSCHE NATIONAL BIBLIOTHEK

Assignment to investigate what a forward-looking subject indexing system might look like

Due to many changes the common rules for subject cataloguing RSWK have to be revised.

The political board for standardisation (STA) charged the expert panel for subject cataloguing to investigate how subject indexing should be done in a forward-looking way.

Expert panel:
- Lively discussions of today‘s conditions

- Intensive study of different systems

DEUTSCHE NATIONAL BIBLIOTHEK

Access to content in the catalogue

Past:

- Large libraries with many items in closed stacks accessible only through the catalogue

- Information about the content only written on the catalogue card

- Subject headings (or numbers from classification scheme) were the only window to the content of a resource (Library catalogues could be used only in the library)

DEUTSCHE NATIONAL BIBLIOTHEK

Access to content in the catalogue

Today:

- Keywords from the catalogue record (online catalogues accessible from all over the world)
- Keywords from digitalised table of contents or from full-text
- Subject Headings: assigned intellectually or by a machine
- Numbers of classification schemes
- Presentation of the catalogue: eg. facettes for searching, drill down menues
- Index lists
- Hyperlinks to other (knowledge organisation) systems
- ....

DEUTSCHE NATIONAL BIBLIOTHEK

Subject indexing with subject headings

The natural language is used

Advantage: metadata are directly accessible in retrieval whereas classification numbers first need to be translated into natural language

Everything mentioned on the slide before can be used for verbal access to the content of a resource

-> this may lead to large lists of results with only minimal relevance

Challenge:

- relevant lists of results
- well structured lists of results

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

- Wikipedia
- Indexing system of a Swiss Television company
- Indexing system of a picture library / archive
- Library of Congress Subject Headings (LCSH)
- Subject indexing by the service provider for public libraries
- Principles of a resource discovery system

Evaluated according to the user tasks („find", „identify", „select", „obtain") of the Functional Requirements for Bibliographic Records

Documentation: Minutes from the Meeting of the Expert Panel, pp. 3-8 (in German only)

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Wikipedia

Not a knowledge organisation system like a classification scheme or a thesaurus but with interesting features for giving guidance to content:

- Hyperlinks (ca. 25 per article)
- System of categories
- Multilingual access: articles in different languages with the same topic are linked together
- Well suited for the semantic web (DBPedia, RDF as standard)

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Indexing system of a Swiss television company

- Indexing according special rules for Swiss television companies
- A detailed abstract is created
- Controlled vocabulary for places and persons
- A database is used
- Target audience: only professionals

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Indexing system of a picture library / archive

– Several categories enabling to find the object (artist, date of manufacture, location, title or name, description of the motive etc.)

– Indexing according to special rules (MIDAS)
  - very well structured: 2200 different fields
  - 18 different controlled vocabularies including a classification scheme
  - temporal search is well developed

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Library of Congress Subject Headings (LCSH)

- Rules and experts for establishing headings (it is an elaborated procedure; it takes time to get headings for new topics)
- Precombination (with elements of precoordination = free floating subdivisions)
- Often easy to understand due to use of conjunctions and prepositions (eg: „Cows on postage stamps“)
- References (synonyms) and access points are not well established

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Subject indexing by the service provider for public libraries

- Four different classification schemes for public libraries are being used
- No subject indexing with subject headings
- Alphabetical lists of „Circles of Interest"; the appellations are regularly revised and adapted by a team of experts

DEUTSCHE NATIONAL BIBLIOTHEK

Review of different systems

Discovery systems

A retrieval tool, not a knowledge organisation system like a classification scheme or a thesaurus, widespread and a basic understanding is useful

- Present different sources of records under one surface
- But the various records from different sources are not mapped on each other, they are not normalised; therefore the results depend strongly on the query
- Although they are more user-friendly than OPACS only experts can fully exploit their potential

DEUTSCHE NATIONAL BIBLIOTHEK

Conclusions of the review

The systems are very different

- *concerning the used indexing methods*
- *concerning the depth of indexing*
- *concerning the target audience*
- *concerning the technical implementation*

Is it possible to find common principles?

Concerning the theoretical concept

Concerning economical aspects

DEUTSCHE NATIONAL BIBLIOTHEK

Conclusions of the review

Common principle concerning the theoretical concept

Well structured data lead to good results

Well structured data = quality indexing (different aspects of the content can be expressed)

Well structured data = basis for a satisfying retrieval (different facettes can be offered; multiple possibilities for presentation)

Well structured data = cannot be generated by a machine alone

DEUTSCHE NATIONAL BIBLIOTHEK

Conclusions of the review

Requirements for well structured data

Binding rules

Consistent application of the rules

If the data are used by other communities or in other applications:
Usage of common standards is highly recommended

An adequate data model

DEUTSCHE NATIONAL BIBLIOTHEK

Conclusions of the review

Common principles concerning economical efficiency

Minimizing production costs by cooperation

- Common rules
- Metadata exchange
- Providing technical interoperability (by using common standards)

Minimizing production costs by using automatical processes

DEUTSCHE NATIONAL BIBLIOTHEK

Summarising the review in 10 guidelines

The Expert Panel summarised the results in 10 guidelines how subject indexing should be done in D-A-CH in a forward-looking way:

Guidelines 1-4: General aspects (aims, scope, target audience, methods)

Guidelines 5-7: Statements concerning a common standard for subject indexing

Guidelines 8-10: Statements concerning the Integrated Authority File (GND) as the most essential tool

Complete guidelines, see Scheven, 2015, slides 24-28 (in German only)

DEUTSCHE NATIONAL BIBLIOTHEK

Thank you for your attention!

Esther Scheven, Deutsche Nationalbibliothek
e.scheven@dnb.de