Thesaurusschulung

Einführung in Thesaurustheorie & (kooperative) Thesauruspflege

Staatliches Museum für Naturkunde Karlsruhe

14. April 2015

Jens M. Lill

Agenda

- Prolog: Motivation zur Nutzung kontrollierten Vokabulars
- Einführung in die Thesaurustheorie
- Einführung in die Thesaurusbearbeitung in imdas pro (Praxisteil)
- Kooperative Thesauruspflege mit imdas pro
  - Organisatorische, inhaltliche, technische Fragen
  - Allgemeine Empfehlungen des BSZ
- Besprechung von konkreten Fragestellungen & Beispielen / Übungen

Kontrolliertes Vokabular

Willkommen in

THESAURUS

Bitte fahren Sie vorsichtig, besonnen, rücksichtsvoll, umsichtig, achtsam, aufmerksam, bedacht durch

Stadt, Ortschaft, Siedlung, Dorf, Gemeinde, Weiler

Wieso? Weshalb? Warum?

- Für eine konsistente Dokumentation
- Für die Qualitätssicherung der Datenerfassung im Museum bzw. Museumsverbund
- Für einen verlässlichen Datenaustauch von Objektdaten
- Für ein verlässliches Retrieval in institutionellen Datenbanken und spartenübergreifenden Kulturgutportalen

Zur Einleitung…

…drei Zitate:

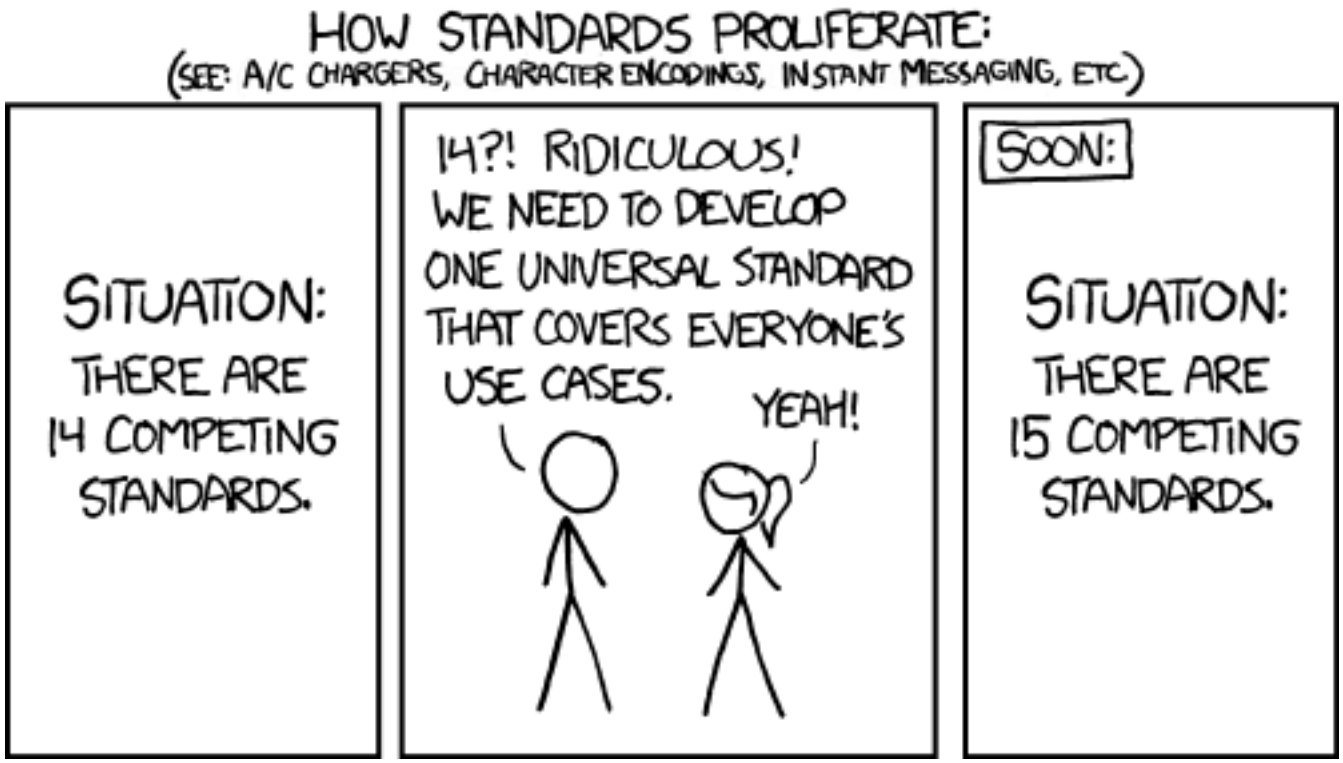
Wenn die Sprache nicht stimmt, dann ist das, was gesagt wird, nicht das, was gemeint ist. (Konfuzius)

[…] kein allgemeines Regelwerk beschränkte den Ideenreichtum und schöpferischen Ordnungsdrang von Generationen wackerer Katalogisierer […] Ohne einen kontrollierten Wortschatz ist ein EDV-Einsatz nicht sinnvoll.
(Datenfeldkatalog zur Grundinventarisation, DMB 1993)

Standards sind wie Zahnbürsten – jeder möchte eine haben, aber keiner die des anderen benutzen!
(Jutta Lindenthal, Thesaurusexpertin, 2010)

Zur Einleitung…

…und ein Cartoon:



Quelle: http://xkcd.com/927/

Objektbezeichnung: Was ist das?

- Ente, Keramikente
- Entenfigur, Figur
- Gefäß mit Deckel in Entenform
- Suppenschüssel mit Deckel (Entenform)
- Deckelterrine/Ente
- Terrine mit Deckel
- Terrine
- *Anas platyrhynchos ceramica*

Wo liegen die Probleme?

- Uneindeutigkeit der natürlichen Sprache führt zu Terminologieproblem
- Bezeichnungen und Begriffe stehen meist nicht in eindeutiger Beziehung zueinander
- **Benennungsvielfalt** (Synonymie)
  z.B.: Terrine = Suppenschüssel
- **Bedeutungsvielfalt** (Homonymie / Polysemie)
  z.B.: Pferd = Tier ; Pferd = Turngerät

→ Vorzugsbegriffe (**Deskriptoren**) bestimmen

Nutzung von Normvokabular

[Sachbegriff **(GND)**] Verwendung: | s |

Sachbegriff: **Terrine**

PPN:	21271869X Zitier
GND-Nummer:	4465433-9 Link zu diesem Datensatz in der GND
Alte Norm-Nr.:	4465433-9 *(in der "**swd**" vor der GND-Migration)*
Frühere Ansetzung:	*in swd:|s|Terrine*
Definition:	Tiefe Schüssel mit Deckel und Griffen
Quelle:	B 1986 unter 1
GND-Systematik:	13.6 *[Kunsthandwerk]* ; 33.2 *[Kochen, Backen, Lebens- und Genussmittel, Küchengerät]*
DDC-Notation:	641.813 ; 642.7

Synonym: Suppenschüssel
Oberbegriff: Essgeschirr *[Oberbegriff allgemein]*

Suche nach Eintrag "**Terrine**" in WIKIPEDIA?

Thesaurustheorie

- Was kennzeichnet einen Thesaurus?
  - Definition und Merkmale nach DIN 1463, Teil 1
    - Begriff vs. Bezeichnung
    - Terminologische Kontrolle
    - Semantische Relationen
- Auswahl und Bildung von Deskriptoren
  - Anforderung an Benennungen (nach DIN 2330)
  - Wortformen
  - Vermeidung von Mehrdeutigkeit
  - Definitionen (Begriffsinhalt und -umfang)
  - Erläuterungen, Quellenhinweise etc.

Thesaurusdefinition nach DIN 1463, Teil 1:

Ein Thesaurus im Bereich der Information und Dokumentation ist eine geordnete Zusammenstellung von **Begriffen** und ihren (vorwiegend natürlichsprachigen) **Bezeichnungen**, die in einem Dokumentationsgebiet zum Indexieren, Speichern und Wiederauffinden dient.

Begriff meint die gedankliche Vorstellung, die sich mit einem Wort verbindet und nicht identisch mit der (individuellen) Bezeichnung ist: z.B. kann die Vorstellung von einer Katze durch die Bezeichnungen *Katze*, *cat*, *Stubentiger, Felis silvestris catus* ausgedrückt werden.

Thesaurusmerkmale nach DIN 1463, Teil 1:

a) Begriffe und Bezeichnungen werden eindeutig aufeinander bezogen (terminologische Kontrolle), indem

- Synonyme möglichst vollständig erfasst werden,
- Homonyme und Polyseme besonders gekennzeichnet werden,
- für jeden Begriff eine Bezeichnung (Vorzugsbenennung, Begriffsnummer oder Notation) festgelegt wird, die den Begriff eindeutig vertritt.

b) Beziehungen zwischen Begriffen (repräsentiert durch ihre Bezeichnungen) werden dargestellt.

Semantische Relationen nach DIN 1463, Teil 1:

- Äquivalenzrelation
- Hierarchierelation
  - Abstraktionsrelation (generische Relation)
  - Bestandsrelation (partitive Relation)
- Assoziationsrelation

Beziehungen eines Deskriptors zu anderen Bezeichnungen zeigen das Begriffsumfeld auf (Verortung im semantischen Gefüge)

→ Begriffsklärung durch das semantische Umfeld

Semantische Relationen nach DIN 1463, Teil 1:

- Äquivalenzrelation (5.2)

Eine Äquivalenzrelation ist die Beziehung zwischen gleichwertigen Bezeichnungen (bedeutungsgleich oder bedeutungsähnlich), die zu einer Äquivalenzklasse zusammengeführt werden.

→ Auflösung der **Benennungsvielfalt** (Synonymie)
→ Schema: Deskriptor BF Synonym
→ Beispiel: Terrine *benutzt für / benutze für* Suppenschüssel

Semantische Relationen nach DIN 1463, Teil 1:

- Hierarchierelation (5.3)

Eine Hierarchierelation liegt vor, wenn zwei Begriffe zueinander in einem Verhältnis der Über- bzw. Unterordnung stehen.

- Abstraktionsrelation (generische Relation)
  Kennzeichen: Oberbegriff / Unterbegriff
- Bestandsrelation (partitive Relation)
  Kennzeichen: Verbandsbegriff / Teilbegriff

Semantische Relationen nach DIN 1463, Teil 1:

- Assoziationsrelation (5.4)

Eine Assoziationsrelation kennzeichnet eine als wichtig erscheinende Beziehung zwischen Begriffen, die weder eindeutig hierarchischer Natur ist, noch als äquivalent angesehen werden kann.

→ Nur angeben, wenn für Indexierung & Retrieval hilfreich
→ Schema: Deskriptor 1 VB Deskriptor 2 (…VB Deskriptor 3)
→ Beispiel: Pyrit *ist ein verwandter Begriff zu* Markasit

Auswahl und Bildung von Deskriptoren

- Anforderung an Benennungen (nach DIN 2330)

Der Zweck von Benennungen ist es, den jeweiligen Begriff innerhalb seines Begriffssystems möglichst genau, knapp und am anerkannten Sprachgebrauch orientiert zu bezeichnen. Benennungen sollen also sprachlich richtig und treffend sein.

→ Genauigkeit durch terminologische Kontrolle

→ Knappheit: Sprechbarkeit, Merkbarkeit, Übersichtlichkeit

→ Sprachgebrauch: Benennungsbildung mi Hilfe von Lexika, Normen, Handbüchern, Wörterbüchern … auch Wikipedia etc.

Auswahl und Bildung von Deskriptoren

- Formale Festlegungen in der Wortform (nach DIN 1463, Teil 1)

→ Wortart: substantivische Form (außer Mehrwort-Deskriptoren)
→ Numerus: bevorzugt im Singular
→ Schreibweise: Photographie vs. Fotografie][evtl. als Synonym
→ Fremdsprachige Bezeichnungen ggf. als Synonyme
→ Abkürzungen und Akronyme ggf. als Synonyme

► Dokumentation (Einleitung/Infoblatt), die Zweck und Aufbau des Thesaurus und das erfasste Gebiet klar darlegt; auch die befolgten Regeln für die Erstellung; statistische Kennzahlen etc.

Auswahl und Bildung von Deskriptoren

- Vermeidung von Mehrdeutigkeit (nach DIN 1463, Teil 1)

→ Kennzeichnung von **Homonymen** (bzw. Polysemen)

Kiefer <Knochen>	Pferd <Tier>
Kiefer <Nadelbaum>	Pferd <Turngerät>

→ **Begriffsdefinitionen** sind erforderlich, wenn Begriffsinhalt und Begriffsumfang eines Deskriptors 1 nicht mit Hilfe der semantischen Relationen innerhalb eines Begriffssystems zur Abgrenzung von Deskriptor 2 ausreichend beschrieben werden kann. (Grundsätze für das Erstellen von Definitionen in DIN 2330 „Begriffe und Benennungen“)

Auswahl und Bildung von Deskriptoren

- Vermeidung von Mehrdeutigkeit (nach DIN 1463, Teil 1)

→ **Erläuterungen** (Scope Notes), auch **redaktioneller Hinweis**: kurze Erklärung des beabsichtigten Gebrauchs eines Deskriptors, um z.B.:

- Anwendungsbereich einzuschränken,
- Abkürzungen und Akronyme zu erklären,
- Änderungen im Gebrauch der Bezeichnung festzuhalten,
- Hinzufügungen / Eliminierungen von Bezeichnungen zu dokumentieren,
- Hinweise der Thesaurusredaktion bei Begriffsansetzung entgegen gängiger Norm etc.

→ **Quellenhinweise** zum Deskriptor bzw. zur Definition

Links im BSZ-Wiki

Bibliotheksservice-Zentrum Baden-Württemberg

Kontrolliertes Vokabular im MusIS-Verbund (Download)

https://wiki.bsz-bw.de/doku.php?id=mare-team:museums-archivsysteme:regelwerke:start

Online-Nachschlagewerke für die Begriffsarbeit

https://wiki.bsz-bw.de/doku.php?id=mare-team:museums-archivsysteme:musis:links:onw

Online-GND (Gemeinsame Normdatei online)

http://swb.bsz-bw.de/DB=2.104/