

Tagging & Erschließung

Bachelor Informationsmanagement
Modul Digitale Bibliothek (SS 2014)

Dr. Jakob Voß

2014-03-24

Problem der Erschließung

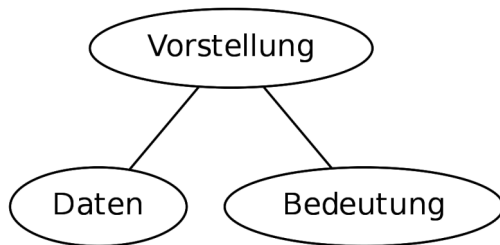
Möglichst *vollständige* und *eindeutige* Beschreibung.

Probleme von Zeichen

- ▶ **Synonyme** (“Bank” = “Sitzbank”)
- ▶ **Homonyme** (“Sitzbank”, “Kreditinstitut” ...)

Lösungsstrategien bei der Erschließung

- ▶ Terminologische Kontrolle / Normdaten
- ▶ Arten von Kontrollierten Vokabularen
 - ▶ Schlagwortliste
 - ▶ Klassifikation
 - ▶ Thesaurus
 - ▶ Ontologie



Zeichen über Zeichen

- ▶ Metadaten zum Dokument
- ▶ Dokument zum Inhalt

Fehlermöglichkeiten

- ▶ mangelnde Terminologearbeit/Datenformate/Programme/...
- ▶ mangelnde Erschließung/Daten/Benutzung/...

...und viele andere Fehler mehr.

Lösungsansätze zur besseren Erschließung

- ▶ Homonymie, Synonymie, Prä/Postkoordination
Terminologische Kontrolle
- ▶ Begriffliche Komplexität
Definitionen und Beispiele!
- ▶ Mangelnde Erschließung
Mehr Mitarbeiter und bessere Werkzeuge

Einführungsbeispiel 1

LibraryThing

- ▶ Gegründet 2005
- ▶ Eigene Bücherlisten/kataloge (10\$/Jahr oder 25\$/Leben)
- ▶ 1.700.000 Mitglieder
- ▶ Inhalte der Seite (sogar Übersetzung) von Mitgliedern
- ▶ Ausgaben, Werke, Rezensionen, Diskussionen etc.
- ▶ Cataloging Parties etc.
- ▶ LibraryThing for Libraries

Einführungsbeispiel 2

Dokumentenserver (z.B. SerWisS unter OPUS 4)

Allgemeine Rollen bei der Erschließung

- ▶ **Autor** erstellt Dokumentes
- ▶ **Archivar** sammelt Dokumente
- ▶ **Dokumentar** beschreibt Dokumente
- ▶ **Leser** findet Dokumente (über Sammlungen & Beschreibungen) und konsumiert Dokumente

Beispiele für Rolleneinteilung

- ▶ Traditionelle Bibliothek

- ▶ Autor:
- ▶ Archivar:
- ▶ Dokumentar:
- ▶ Leser:

- ▶ Dokumentenserver

- ▶ Autor:
- ▶ Archivar:
- ▶ Dokumentar:
- ▶ Leser:

- ▶ LibraryThing

- ▶ Autor:
- ▶ Archivar:
- ▶ Dokumentar:
- ▶ Leser:

Was hat das mit digitalen Bibliotheken zu tun?

- ▶ Digitale Bibliotheken beschäftigen sich mit Daten
 - ▶ Erschließungsdaten (Metadaten)
 - ▶ Dokumente
- ▶ Einteilung im digitalen flexibel
 - ▶ Metadaten / Dokumente
 - ▶ Autor / Leser / Archivar / Dokumentar
- ▶ Automatische Prozesse in allen Bereichen

Social Cataloging

- ▶ Erschließung durch Benutzer
- ▶ Beispiele:
 - ▶ LibraryThing, Comics.org, IMBD, ISFDB...
 - ▶ BibSonomy, Mendeley, Zotero...
 - ▶ Wikidata
 - ▶ ...
- ▶ Social Bookmarking

Beispiel 3: BibSonomy

Bag-Model vs. Set-Model

- ▶ Bag: jeder hat seins
- ▶ Set: alle an einem

Social Cataloging *light* in Bibliotheken

- ▶ Verbundkatalogisierung (*Set oder Bag?*)

Social Cataloging in Bibliotheken

- ▶ Export-Möglichkeiten in Social-Cataloging-Plattformen
- ▶ Kaum erfolgreiche Beteiligung am Katalogisieren
 - ▶ Technische Probleme
 - ▶ Mangelnde Usability
 - ▶ Fehlende Motivation
 - ▶ Bestand statt Relevanz
 - ▶ ...

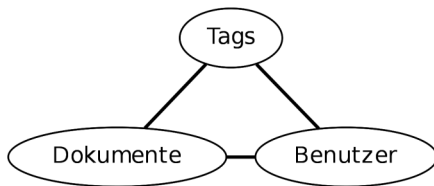
Wer erschließt stattdessen?!

Social Tagging

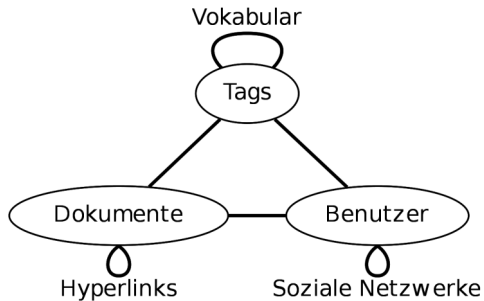
Beispiele

- ▶ StackExchange
- ▶ <http://www.flickr.com/photos/tags/>

Social Tagging-Modell: Tripartiter Graph



Social Tagging-Modell



Aggregation von Social Tagging-Daten

- ▶ 3-Dimensionale Matrix (Tags, Benutzer, Dokumente)
- ▶ Jeder Tagging-Vorgang ist eine Koordinate
- ▶ etablierte Verfahren im Data Warehousing

Aggregation von Social Tagging-Daten

- ▶ Aggregation über
 - ▶ alle Benutzer (Tags / Dokumente)
 - ▶ alle Dokumente (Tags / Benutzer)
 - ▶ alle Tags (Benutzer / Dokumente)
 - ▶ Benutzer & Tags = Persönliche Liste
 - ▶ Benutzer & Dokumente = Persönliche Themen
 - ▶ ...

Weitere Auswertungen möglich

- ▶ Ranking über Verlinkung (FolkRank)
- ▶ Synonyme über Kozitation
(Tags and gemeinsamen Dokumente)
- ▶ Hyponyme über Bibliographische Kopplung
(gemeinsame Tags an anderen Dokumenten)

Arten von Tagging-Systeme

- ▶ Tagging \neq Tagging
- ▶ Verschiedene Aspekte der Unterscheidung
- ▶ Akzeptanz & Nutzen je nach Umsetzung

Arten von Tagging-Systemen

resource sources woher stammen die Dokumente?

(neu erstellt, frei gefunden, vorhandene Sammlung...)

resource representation wie liegen Dokumente beim Tagging vor?

resource connectivity Hierarchien & Hyperlinks

Arten von Tagging-Systemen

tagging-rights wer darf Tags vergeben?

(Autor, Bibliothekar, Benutzer, Rollen...)

tag aggregation set-Model (ein Tag-Set pro Dokument)

bag-Model (jedem sein eigenes Tag-Set)

Arten von Tagging-Systemen

tagging feedback Rückmeldung bei der Vergabe von Tags?
automatic tagging (Teil)automatisches Tagging?

Arten von Tagging-Systemen

tag control Kontrolliertes Vokabular?

(Synonyme, Umbenennung, Verknüpfung...)

tag connectivity flach, Hierarchie, Thesaurus, Facetten...

Arten von Tagging-Systemen

- ▶ Kriterien ähnlich bei allen Systemen zur Schlagwortvergabe
- ▶ Kriterien ähnlich bei allen Systemen zur Datenerfassung