



Universität Augsburg
Universitätsbibliothek

Herausforderungen der Inhaltserschließung mit GND- Schlagwörtern nach RSWK

Einblicke in die Arbeit der AG SE – Teil1

Martin Völkl

25.09.2025

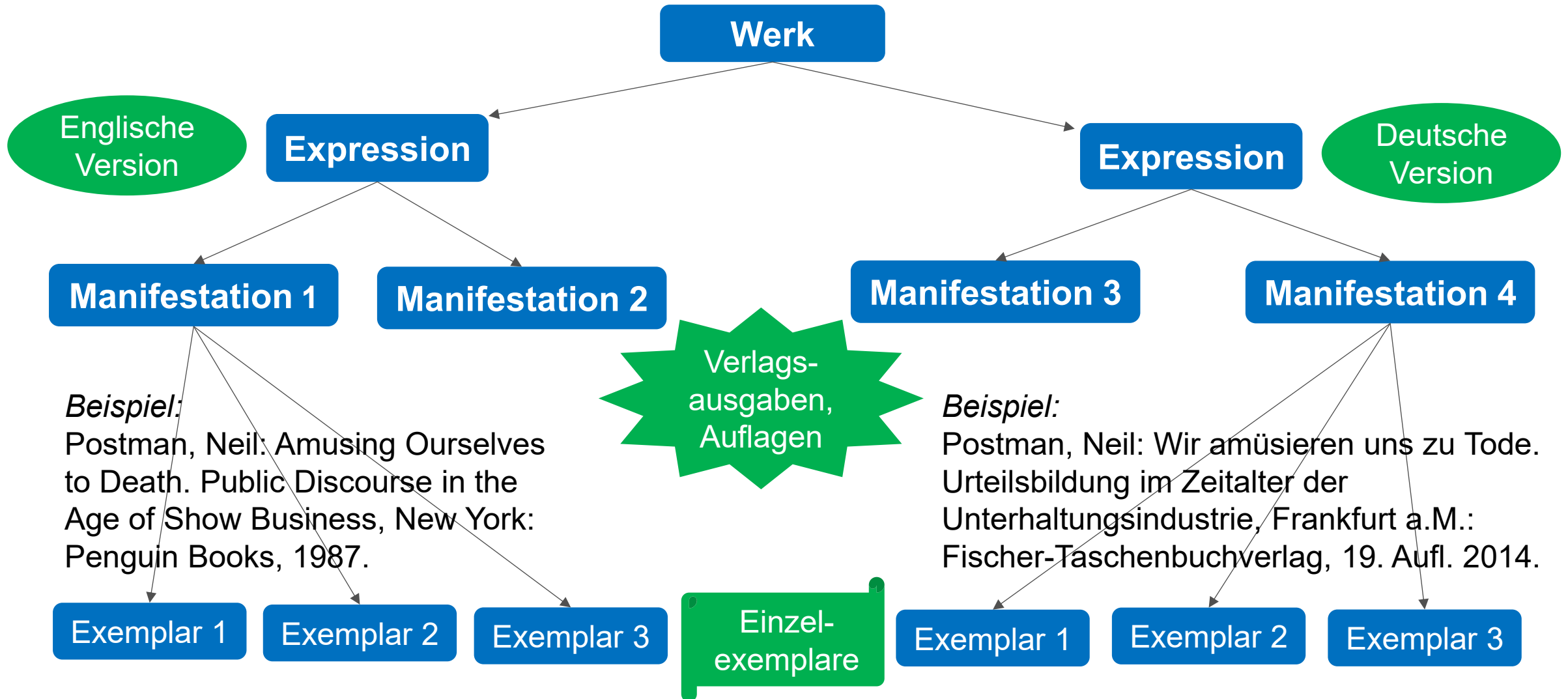
Agenda

- 1** Input: Intellektuelle versus maschinelle Inhaltserschließung? Das Culturegraph-Projekt der AG SE
- 2** Output: Fehlertypen bei der Kataloganreicherung mit Schlagwortfolgen und Formangaben
- 3** Outcome: Tipps für die intellektuelle Inhaltserschließung mit GND-Schlagwörtern nach RSWK
 - 3a** Tipp 1: Verschlagwortung von Expressionsmerkmalen
 - 3b** Tipp 2: Verwendungshinweise lesen
 - 3c** Tipp 3: Übersetzungen
 - 3d** Tipp 4: Mehrfach dieselbe Formangabe
 - 3e** Tipp 5: Gleichzeitig Autobiografie und Biografie

1. Input: Intellektuelle versus maschinelle Inhaltsererschließung? Das Culturegraph-Projekt der AG SE

- Goldstandard der verbalen Inhaltsererschließung im DACH-Raum: Intellektuelle Erschließung einer vorliegenden Ressource (Autopsieprinzip) mit GND-Schlagwörtern nach RSWK
- Weitere Formen der verbalen Inhaltsererschließung im DACH-Raum:
 - Maschinelle Inhaltsererschließung (⇒ automatisierte Zuweisung von inhaltserschließenden Deskriptoren aus der GND an Ressourcen durch lernende Systeme / künstliche Intelligenz; Bsp. Maschinelle Erschließung der DNB oder der ZBW)
 - Maschinelle Unterstützung der intellektuellen Inhaltsererschließung, z.B. durch das *Unterstützungstool für Fachreferatsarbeit* der UB Mannheim, oder durch den *Digitalen Assistenten 3*)
 - Maschinelle Übertragung der intellektuell erstellten, inhaltserschließenden Metadaten (⇒ Formangaben, GND-Schlagwortfolgen nach RSWK, Notationen verschiedener KOS) von einer Manifestation eines Werks auf andere Manifestationen desselben Werks durch Culturegraph
 - ⇒ Werkclusterverfahren, in dem alle Manifestationen eines Werks gebündelt und die inhaltserschließenden Metadaten einer Manifestation auf andere übertragen werden
 - ⇒ Grundlage: Im Gegensatz zur Formalerschließung wird im Rahmen der Inhaltsererschließung in der Regel nicht die vorliegende Manifestation, sondern das betreffende Werk erschlossen.

1. Input: Intellektuelle versus maschinelle Inhaltsererschließung? Das Culturegraph-Projekt der AG SE



1. Input: Intellektuelle versus maschinelle Inhaltsererschließung? Das Culturegraph-Projekt der AG SE

| | |
|------|---|
| 902 | s Massenmedien 9 (DE-588)4037877-9 |
| 902 | s Kritik 9 (DE-588)4033229-9 |
| 903 | a 21 |
| 907 | g USA 9 (DE-588)4078704-7 |
| 907 | s Massenmedien 9 (DE-588)4037877-9 |
| 907 | s Öffentliche Meinung 9 (DE-588)4043152-4 |
| 908 | a 213 a 312 |
| 909b | a 1@@acgwrk@@d20201028@@qDE-101@@ |
| 912 | s Telekommunikation 9 (DE-588)4059360-5 |
| 912 | s Kultur 9 (DE-588)4125698-0 |
| 912 | s Zukunftserwartung 9 (DE-588)4191125-8 |
| 913 | a 213 a 312 |
| 914b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 917 | s Massenmedien 9 (DE-588)4037877-9 |
| 917 | s Gesellschaft 9 (DE-588)4020588-5 |
| 918 | a 21 |
| 919b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 922 | s Fernsehen 9 (DE-588)4016825-6 |
| 922 | s Gesellschaft 9 (DE-588)4020588-5 |
| 923 | a 21 |
| 924b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 927 | s Urteilen 9 (DE-588)4078671-7 |
| 927 | s Unterhaltungsindustrie 9 (DE-588)4186999-0 |
| 928 | a 21 |
| 929b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |

Titelsätze in ALEPH nach der Culturegraph-Anreicherung (am Beispiel der beiden o.g. Manifestationen von Neil Postmans Buch „Amusing Ourselves to Death“ / „Wir amüsieren uns zu Tode“): links die englische, rechts die deutsche Ausgabe.

Die englische Ausgabe von 1987 hatte bereits eine Schlagwortfolge im Feld 902.

Angereichert wurden hier weitere Schlagwortfolgen (Felder 907, 912, 917, 922 und 927), die jeweils mit einer Provenienzanabe versehen wurden („acgwrk“).

| | |
|------|---|
| 902 | g USA 9 (DE-588)4078704-7 |
| 902 | s Massenmedien 9 (DE-588)4037877-9 |
| 902 | s Öffentliche Meinung 9 (DE-588)4043152-4 |
| 903 | a 213 a 312 |
| 907 | s Telekommunikation 9 (DE-588)4059360-5 |
| 907 | s Kultur 9 (DE-588)4125698-0 |
| 907 | s Zukunftserwartung 9 (DE-588)4191125-8 |
| 908 | a 213 a 312 |
| 909b | a 1@@acgwrk@@d20201028@@qDE-101@@ |
| 912 | s Massenmedien 9 (DE-588)4037877-9 |
| 912 | s Kritik 9 (DE-588)4033229-9 |
| 913 | a 21 |
| 914b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 917 | s Massenmedien 9 (DE-588)4037877-9 |
| 917 | s Gesellschaft 9 (DE-588)4020588-5 |
| 918 | a 21 |
| 919b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 922 | s Fernsehen 9 (DE-588)4016825-6 |
| 922 | s Gesellschaft 9 (DE-588)4020588-5 |
| 923 | a 21 |
| 924b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 927 | s Urteilen 9 (DE-588)4078671-7 |
| 927 | s Unterhaltungsindustrie 9 (DE-588)4186999-0 |
| 928 | a 21 |
| 929b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |

Die deutsche Ausgabe hatte ebenfalls eine Vorerschließung mit einer Schlagwortfolge im Feld 902. Diese wurde in der englischen Ausgabe in Feld 907 angereichert.

Die anderen Schlagwortfolgen wurden angereichert und mit einer Provenienzanabe („acgwrk“) versehen.

2. Output: Fehlertypen bei der Kataloganreicherung mit Schlagwortfolgen und Formangaben

- Ergebnisse der ersten Anreicherungsphase:
 - Insgesamt wurden ca. 1,7 Millionen Print-Titeldatensätze und knapp 470.000 Titelsätze von eBooks mit inhaltserschließenden Metadaten angereichert.
 - Bei knapp 900.000 Titeldatensätzen wurden Formangaben zur Art des Inhalts angereichert.
 - 1.004.177 angereicherte Titelsätze wiesen vorher keine inhaltliche Erschließung mit Einzelschlagwörtern oder Schlagwortfolgen auf.
- Die Analyse fehlerhafter Anreicherungen ergab **drei verschiedene Fehlertypen:**
 - **Maschinell** ⇒ fehlerhafte Bündelung ⇒ in einem Werkcluster sind Manifestationen unterschiedlicher Werke enthalten (dies war v.a. bei großen Clustern der Fall)
 - **Strukturell** ⇒ regelwerkskonforme Erschließung führt trotz korrekter Werkcluster zu inhaltlich fehlerhaften Anreicherungen
 - **Intellektuell** ⇒ nicht regelwerkskonforme Verwendung von Formangaben und Schlagwörtern, die infolge des Culturegraph-Projekts auf andere Manifestationen desselben Werks übertragen wurden



2. Output: Fehlertypen bei der Kataloganreicherung mit Schlagwortfolgen und Formangaben

- Insgesamt wurden bisher 13 Bereinigungsaktionen durchgeführt (⇒ Identifikation der betreffenden Titeldatensätze über die Provenienzkennung der angereicherten Formangaben und/oder Schlagwortfolgen)
- Zusätzlich wurde in ASEQ-Feld 078n die Möglichkeit geschaffen, intellektuell vorgenommene Korrekturen im betreffenden Titeldatensatz zu vermerken (⇒ „cgwrk-korr“).
- Beispiel für die **Rücknahme struktureller Fehler** der Culturegraph-Anreicherung **von bestimmten Formangaben**
 - Ein historischer Quellentext kann im B3Kat sowohl als Inkunabel, als auch als historisch-kritische Edition vorliegen, die gleichzeitig auch eine Hochschulschrift sein kann.
 - Problem: Alle genannten Fälle sind Manifestationen desselben Werks, sind also im selben Werkcluster enthalten.
 - Titeldatensätze dieses Bündels, die noch keine Formangabe aufwiesen, erhielten fälschlicherweise alle drei Formangaben Inkunabel, Quelle und Hochschulschrift.

2. Output: Fehlertypen bei der Kataloganreicherung mit Schlagwortfolgen und Formangaben

- Beispiel für die **Rücknahme struktureller Fehler** der Culturegraph-Anreicherung **von bestimmten Schlagwortfolgen**
 - Befinden sich in dem Bündel eines Werks, das eine historische Entwicklung bis zur Gegenwart schildert, verschiedene inhaltlich erweiterte Auflagen, dann bildet sich das Erscheinungsdatum der jeweiligen Auflage i.d.R. auch in der Schlagwortfolge ab.
 - Beispiel:

Castles, Stephen / Miller, Mark J.: The age of migration. International population movements in the modern world, 4. Auflage, Basingstoke [u.a.] 2009.

⇒ Im B3Kat ist diese Manifestation nach der Culturegraph-Anreicherung mehrfach erschlossen mit derselben Schlagwortfolge \$s Internationale Migration / \$z Geschichte, wobei sich diese Schlagwortfolgen nur in der zeitlichen Erstreckung unterscheiden: 1945-1990, oder 1945-2001, oder 1945-2008.
 - Erschlossen ist hier (ausnahmsweise) nicht das Werk, sondern die jeweilige Expression
 - ⇒ Diese Erschließung entspricht den RSWK, ist also regelwerkskonform!

2. Output: Fehlertypen bei der Kataloganreicherung mit Schlagwortfolgen und Formangaben

| | |
|------|---|
| 927 | s Internationale Migration 9 (DE-588)4162051-3 |
| 927 | z Geschichte 1945-1990 ← |
| 929b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 932 | s Internationale Migration 9 (DE-588)4162051-3 |
| 932 | z Geschichte 1945-2008 ← |
| 934b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 937 | s Migration 9 (DE-588)4120730-0 |
| 937 | z Geschichte |
| 939b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 942 | s Internationale Migration 9 (DE-588)4162051-3 |
| 942 | z Geschichte 1945-1996 ← |
| 944b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |
| 947 | s Internationale Migration 9 (DE-588)4162051-3 |
| 947 | z Geschichte 1945-1993 ← |
| 949b | a 1@@acgwrk@@d20201028@@qDE-101@@uhttps://d-nb.info/provenance/plan#cgwrk |

In diesem Beispiel (bei ansonsten gleicher Schlagwortfolge) vier verschiedene zeitliche Erstreckungen beim Zeitschlagwort „Geschichte“, die sich ursprünglich auf unterschiedliche Auflagen bezogen.

Tipp: Bei der Übernahme von Zeitschlagwörtern aus Voraufgaben: Überprüfung der zeitlichen Erstreckung!

Vielen Dank
für Ihre Aufmerksamkeit



Martin Vökl
Universitätsbibliothek
Universität Augsburg
martin.voelkl@bibliothek.uni-augsburg.de
www.uni-augsburg.de