

Abbildung der Werkebene nach RDA Stand der Überlegungen in der AG RDA

Inhalt

- (1) Policy der AG RDA zu Werk und Expression
- (2) Werke und Teil-Ganzes Beziehungen
- (3) Clusterverfahren für Werke
- (4) Laufende und geplante Aktivitäten

(1)

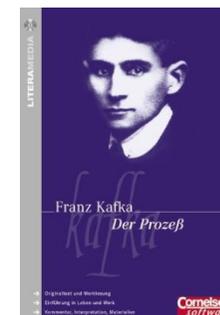
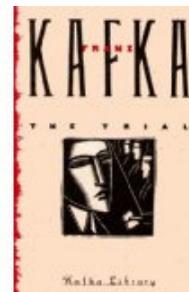
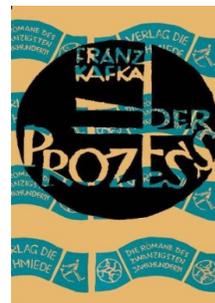
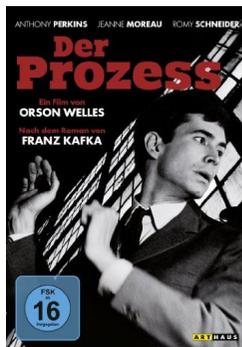
Zusammenführen, was zusammengehört



<http://d-nb.info/gnd/4099250-0>
Kafka, Franz
Der Prozeß



RDAs use of FRBR entities makes it possible ... clustering information about the same title together



(1)

Ergebnis der Diskussion in der AG RDA Stand März 2013

- Gemäß Implementation Scenario 2 wird eine Composite Description mit beschreibenden Elementen für die FRBR Entitäten Werk, Expression, Manifestation angelegt.
- Die FRBR-Entität Expression wird zum Zeitpunkt der Einführung der RDA nicht als eigenständige Entität repräsentiert.
- Unter bestimmten Bedingungen sollen intellektuell Normdatensätze für Werke in der GND erfasst bzw. vorhandene Werksätze nachgenutzt werden.
- Das automatische Clustern von Manifestationen eines Werkes wird angestrebt, jedoch wird eine Realisierung im laufenden Projekt nur von der DNB, dem OBV und dem DLA Marbach verfolgt. (Werksatz wird angelegt, wenn ≥ 2 Expressionen/Manifestationen zu demselben Werk vorliegen)
- Zunächst wurde die Entscheidung zurückgestellt, ob basierend auf dem Clusterverfahren Normdatensätze der GND generiert werden sollen.

(2)

Die Wirklichkeit ist kompliziert

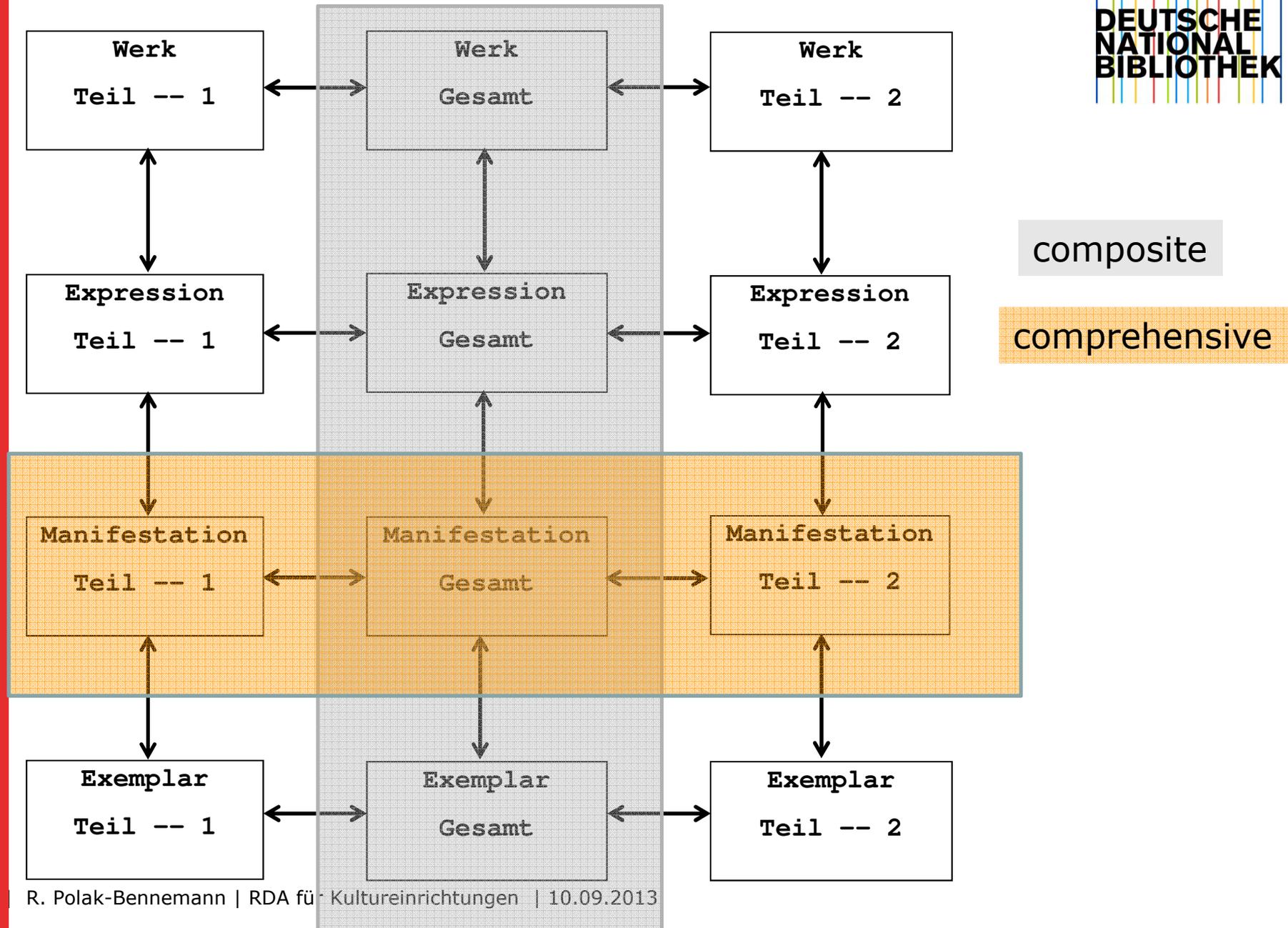
The World`s Thickest Book: Agatha Christie: „The Complete Miss Marple“



© HarperCollins



(2) FRBR-Entitäten und Teil-Ganzes-Beziehungen



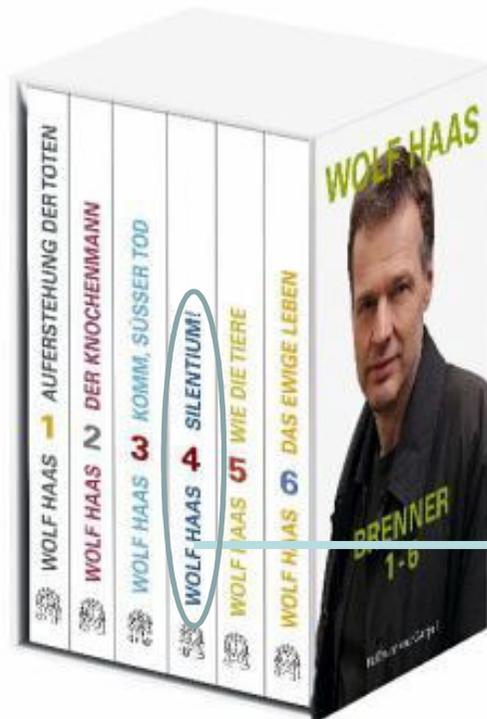
(2)

Gesamtwerk mit Stücktiteln

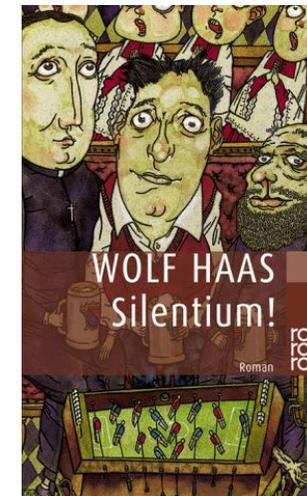
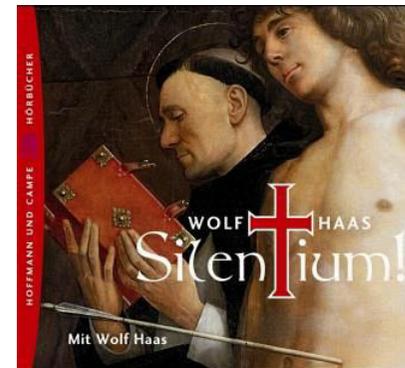
Arbeitshypothese:

Das Gesamtwerk ist ein Werk im Sinne von RDA.

Ein Stücktitel ist ein Werk im Sinne von RDA.



Expressionen
Manifestationen →



(2)

Gesamtwerk mit Bandaufführungen



Arbeitshypothese:

Das Gesamtwerk ist ein Werk im Sinne von RDA.

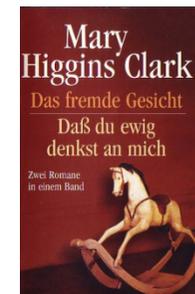
Ein unselbständiger Bandsatz ist kein eigenes Werke im Sinne von RDA.

Compilations - Typen

- Compilations mit übergeordnetem Titel
 - „enthaltene Werke“
 - Beiträge in Sammelwerken
 - ...



- Compilations ohne übergeordnetem Titel
 - „beigefügte Werke“
 - ...



Arbeitshypothese:

Eine Compilation ist ein Werk im Sinne von RDA. Die in der Compilation enthaltenen Teile sind Werke im Sinne von RDA.

(3)

Erprobung von Algorithmen zum Werkclustering



- Plan
 - 3 Testphasen mit jeweils Optimierung des Algorithmus, Testlauf und Bewertung mit ‚Alt-‘Daten im MARC 21 Format aus der DNB, des OBV und des DLA
 - Eine weitere Testphase soll mit RDA-Daten durchgeführt werden (ca. 2. Quartal 2014).

- Matchkey Phase 1
 - 2 Elemente: Titelement und, falls vorhanden, Elemente von Personen, Körperschaften und Kongressen
 - Normalisierung : Auflösung von Umlauten, Diakritika, Artikel, Abkürzungen, Zahlenwörter etc.

- Testset Phase 1
 - 969 Titeldatensätze aus dem DNB-Bestand
 - umfasst Titel, die zu einem Werk gehören, Titel, die nicht zu demselben Werk gehören und zufällig ausgewählte Titel

(3)

Beispiel: Bündel zum gleichen Werk

Anzahl Sätze	Matchkey	Bemerkung
78	lestroismousquetaires dumasalexandre	EST/Autor
11	die3musketiere dumasalexandre	Titel/Autor
6	lestroismousquetaires dumasalexandre <u>re</u> pe	EST/Autor (Text)
1	lestroismousquetaires dumasalexandre	EST/Autor; Schreibfehler i
1	lestroism <u>u</u> squetaires dumasalexandre	EST/Autor; Schreibfehler u

Finden derzeit noch nicht zusammen (diverse Gründe) !



2. Matchkey: die3musketiere dumasalexandre

Normalisierung korrigieren

Ähnlichkeitsvergleich

GND einbeziehen; Textform (= Variante Namensform) in bevorzugte Benennung ersetzen

(3)

Beispiel: Bündel zum gleichen Werk - Tonträger

Anzahl Sätze	Matchkey	Bemerkung
1	die3musketiere dumasalexandrekrejewalterassheuerbirgitta	Titel/Beteiligte; Neu erzählt von.... Eigenes Werk?
1	die3musketieredumasalexandreprimusbodo	Titel/Beteiligte; Gekürzte Lesung
2	lestroismousquetaires dumasalexandreprimusbodo	EST/Beteiligte Gekürzte Lesung

Finden derzeit noch nicht
zusammen (RAK-Regel für
Tonträger) !



2. Matchkey:
die3musketieredumasalexandreprimusbodo
Weitere Optimierungen?
Bei Tonträger nur die erste beteiligte Person?
Das bedeutet: Unschärfe, z. B. sind dann auch Hörspiele im gleichen
Cluster.

(3)

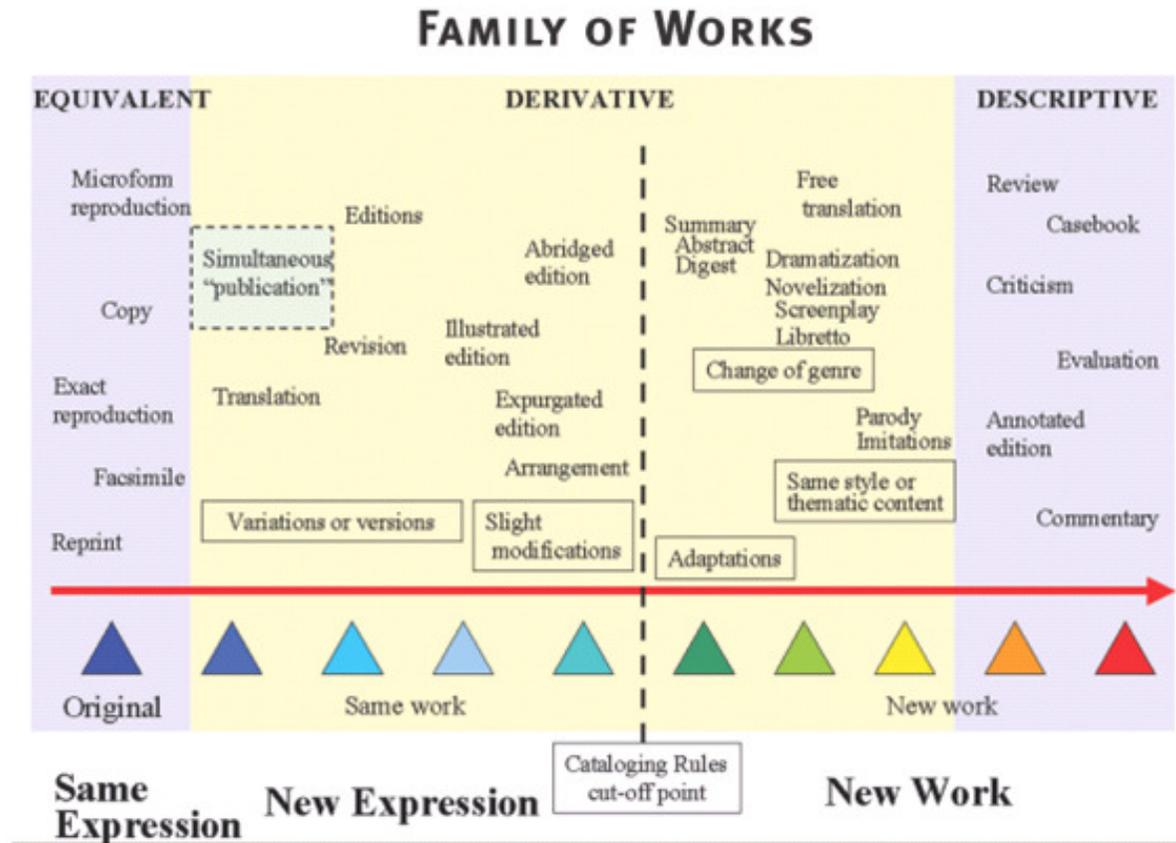
Wann ist ein Werk ein Werk? (# 1)

- Ist die Reihenfolge der Verfasser relevant für die Entscheidung, ob es sich um dasselbe Werk oder ein neues Werk handelt?
- Gehören zwei Manifestationen, die denselben Titel und 1. Verfasser haben, zu demselben Werk, falls einer der Titel einen zusätzlichen Verfasser (700 \$4 aut) aufweist?
- „... neu erzählt...“ - handelt es sich dabei um ein neues Werk?
-

(3)

Wann ist ein Werk ein Werk? (#2)

- in welchen Fällen handelt es sich um ein neues Werk?
- in welchen Fällen gehören Expression /Manifestation zu demselben Werk?
- in welchen Fällen stehen Werke zueinander in Beziehung?



Graphik: Barbara Tillett, Family of Works, S. 4
<http://www.loc.gov/cds/downloads/FRBR.PDF>

Relationships in the Organization of Knowledge, edited by Carol A. Bean and Rebecca Green, 2001, p. 23, "Bibliographic Relationships" by Barbara B. Tillett, Figure 2, © 2001 Kluwer Academic Publishers Boston, with kind permission of Kluwer Academic Publishers.

(4)

Laufende und geplante Aktivitäten (# 1)

- Erfassungspraxis für Werksätze der Musik prüfen → UAG Musik
- Erfassungspraxis für Werksätze prüfen → UAG GND (ab 2014)
- Werke bei Fortlaufenden Sammelwerken → UAG FS (Sommer 2013)
- Entscheidungen für Werke und Teil-Ganzes Beziehungen treffen → Sommer/Herbst 2013
- Evaluierung Clusterverfahren → Sommer/Herbst 2013 / 2. Quartal 2014

(4)

Laufende und geplante Aktivitäten (#2)

- Schärfung des Werkbegriffs → Herbst 2013
- Entscheidung über automatisch generierte Werksätze in der GND treffen → bis 3. Quartal 2014
- Workflow (Datenfluss, Verlinkung) → spätestens ab 3. Quartal 2014

Danke!

(R. Polak@dnb.de)