



hbz

Wissen. Information. Innovation.

Matching von Titeldaten - Fragen und Ansätze

DINI AG KIM Workshop, Mannheim

11. April 2012

Adrian Pohl <pohl@hbz-nrw.de>

Agenda

- Grundbegriffe
- Identifikation bibliographischer Ressourcen
- Zwischenfazit
- Matching & Kontext

Grundbegriffe

(Bibliographische) Ressource

- Dinge/Entitäten/Objekte, die in Katalogen, Bibliographien etc. beschrieben werden
- Exemplar, Auflage, Manifestation, Werk ...
- Monographie, Aufsatz, Journal, Mehrbändiges Werk, Band etc.
- Print, elektronisch ...

Was wird katalogisiert?

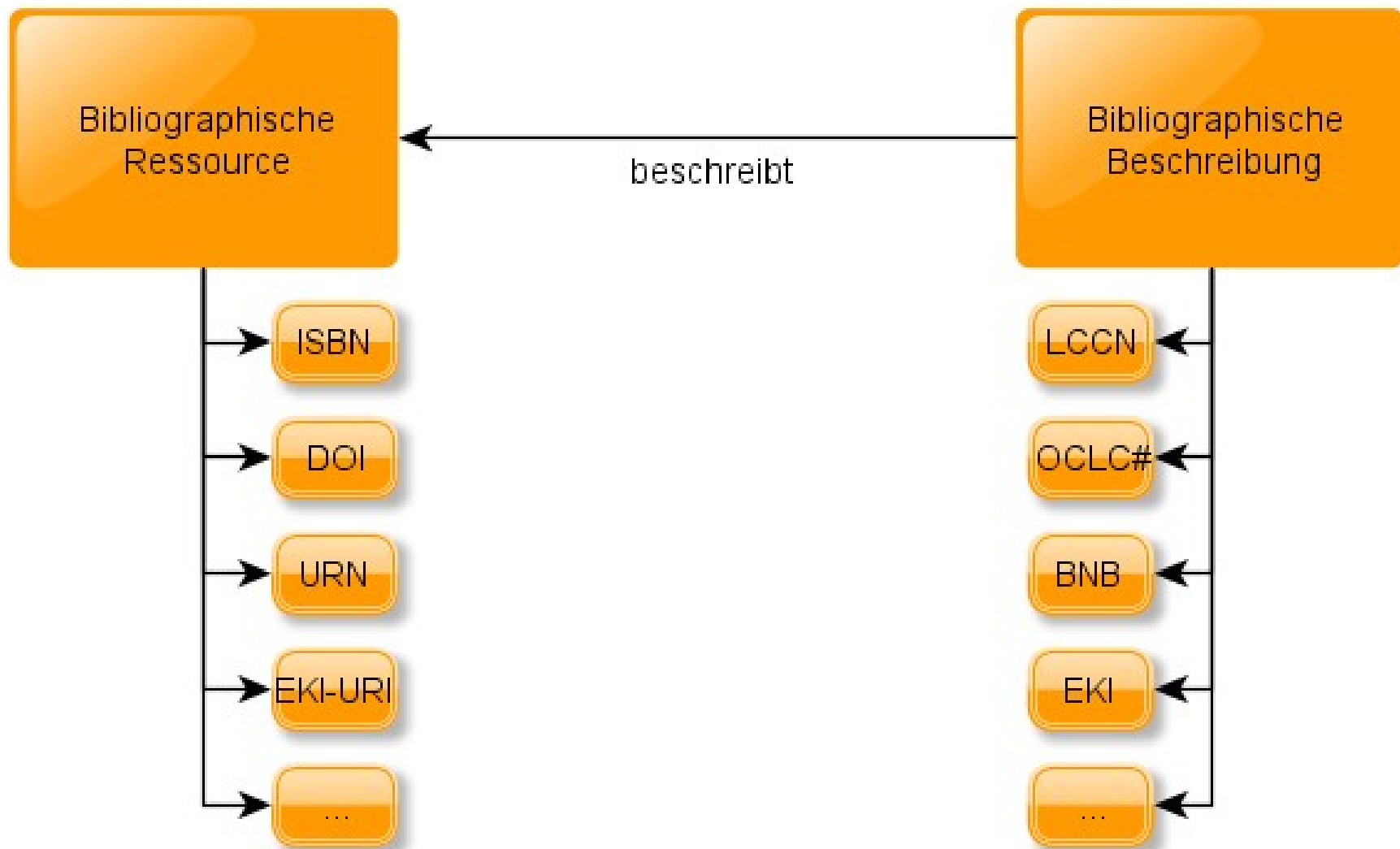
Menge von Kopien /
Typ

Exemplar /
Instantiierung eines Typen

Menge von Mengen von Kopien /
Menge von Typen

(Bibliographische) Beschreibung

- Beschreibungen von Dingen (von Exemplaren, Typen, Mengen von Typen)
- Das, was in RDF repräsentiert wird



Identifikation bibliographischer Ressourcen

Feuilletonistisch

das erste Hauptwerk
des österreichischen
Philosophen Ludwig
Wittgenstein

Bibliographisch

Ludwig Wittgenstein: Tractatus logico-philosophicus, Logisch-philosophische Abhandlung. Suhrkamp, Frankfurt am Main 2003. ISBN 3-518-10012-2

Mittels Identifikatoren

ISBN 3-518-10012-2

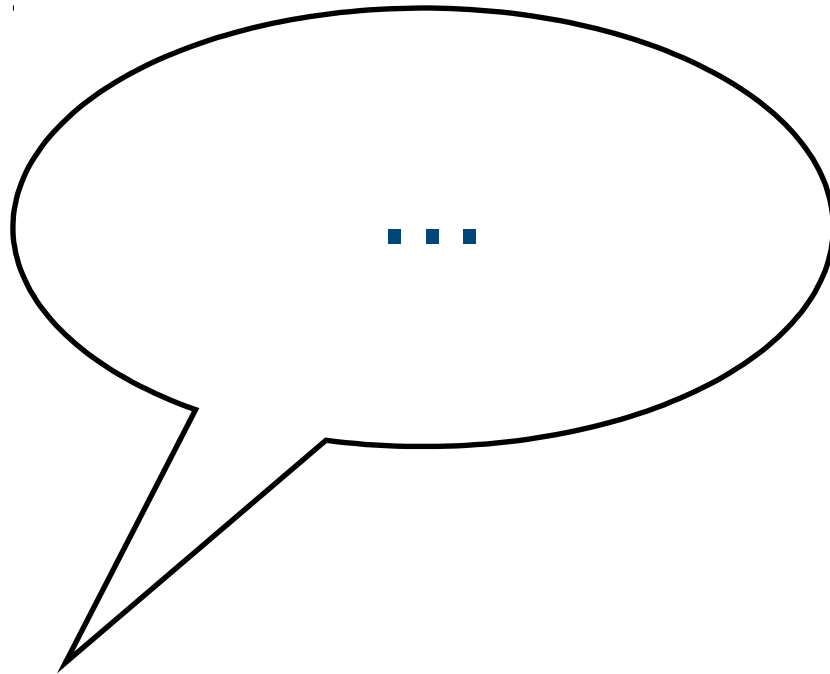
Das Buch, dessen
WorldCat-Eintrag die
OCLC-Nummer
'174202788' hat.

<http://openlibrary.org/works/OL15680951W>

Durch Hinweisen

"Das dunkelblaue Buch
dort auf dem Tisch."

<http://www.gutenberg.org/files/5740/5740-pdf.pdf>



Was wird hier identifiziert?

- Ein Exemplar
- So etwas wie ein FRBR-Werk
- mehrere Druckauflagen eines Taschenbuchs
- eine konkrete Druckauflage
- Eine Übersetzung + Vorwort als E-Book
- ...

Zwischenfazit

- Es gibt nicht *die* Art der Identifikation bibliographischer Ressourcen, sondern unterschiedliche Praktiken der Identifikation.
- Die richtige Art der Identifikation ist einer Frage der Bedürfnisse/des Kontexts.
- Eine Identifikationspraxis sollte sich an den Bedürfnissen ihrer Anwender orientieren.

Um sinnvolle (Un)Gleichheitsaussagen über zwei Dinge machen zu können, müssen diese zur selben Kategorie gehören, d.h. derselben Identifikationspraxis folgen.

Unsinn:

das erste Hauptwerk des Philosophen Ludwig Wittgenstein

=

Ludwig Wittgenstein: Tractatus logico-philosophicus, Logisch-philosophische Abhandlung. Suhrkamp, Frankfurt am Main 2003. ISBN 3-518-10012-2

=

ISBN 3-518-10012-2

=

Das dunkelblaue Buch dort auf dem Tisch

Matching & Kontext

Praxisfragen für Matcherinnen & Matcher

- Was wird identifiziert? (Beschreibung oder Ressource)
- Welche Identifikationspraxis liegt den zu matchenden Ressourcen zugrunde?
- Welches sind identifizierende und unterscheidende Merkmale bei meinem Matching?

Kontext dokumentieren

- Bei der ID-Vergabe wie auch beim Matching ist es sinnvoll, den Kontext zu dokumentieren.
- Die Bundle-Ontologie bietet die Möglichkeit, den Kontext für die Bündelung ähnlicher Ressourcen in RDF zu dokumentieren.
- `<http://purl.org/net/bundle#>`



Eintrag match-oclc\$key:10009258

Daten

cg:type	bundle
cg:contains	BSZ-009000402
	BVB-BV000256321
	BVB-BV000912722
	GBV-031104118
	GBV-148302475
	HEB-056874464
	HEB-235822388
cg:alg	match-oclc

Danke.

Fragen?

Gerne jetzt oder auch später an
semweb@hbz-nrw.de

Lizenz



Diese Folien stehen unter folgender CC-Lizenz:

<http://creativecommons.org/licenses/by/3.0/de/>