

Linked Open Data in Bibliotheken, Archiven & Museen

Jakob Voß

Verbundzentrale des GBV (VZG)

Infocamp, Chur, 2014-10-10



Häh?

- ▶ Alles hängt irgendwie mit allem zusammen
- ▶ Hier: **Daten** die mit anderen Daten zusammenhängen
- ▶ Linked Open Data

Daten die mit anderen Daten zusammenhängen

- ▶ Verschiedene Daten beschreiben die gleiche Dinge
- ▶ Zum Beispiele **die gleiche Person**
 - ▶ Person als Autor in Katalogdatenbank
 - ▶ Objekte der Person im Bestandsverzeichnis eines Museums
 - ▶ Wikipedia-Artikel über die Person
 - ▶ ...

Linked Open Data ist überall relevant, wo in verschiedenen Einrichtungen Daten über gleiche Dinge verwaltet werden

Old School Linked (Open) Data: authority files

Normdaten Personenverzeichnis, Klassifikation, Thesaurus. . .
Kontrollierte Vokabulare statt Hymonyme und Synonyme
Identifier Notation, ID-Nummer. . .

Grundidee: Eindeutige Referenzierbarkeit

“things, not strings”

things, not strings – ein Beispiel

1. “Autor: Karl Marx”
“Autor: Karl Marx (Künstler)”
2. “Autor:” <http://d-nb.info/gnd/118578545>
3. Eigenschaft: <http://purl.org/dc/terms/creator>
Gegenstand: <http://d-nb.info/gnd/118578545>

⇒ Linked Open Data

Linked Open Data

1. **Data**

Daten in RDF

HTTP-URLs als Identifier

2. **Open**

abrufbar per HTTP-URLs

3. **Linked**

mit Links zu anderen HTTP-URLs

Beispiel: <http://d-nb.info/gnd/118578545>

GND	
Link zu diesem Datensatz	http://d-nb.info/gnd/118578545
Person	Marx, Karl
Quelle	B 1996; Internet
Zeit	Lebensdaten: 1929-2008
Land	Deutschland (XA-DE)
Geografischer Bezug	Geburtsort: Köln Sterbeort: Köln
Beruf(e)	Künstler
Weitere Angaben	Dt. Maler; 1946- Studium der freien Malerei an den Kölner Werkschulen; 1959- Dozent an den Kölner Werkschulen und später Professor für Malerei; 1975- Dekan des Fachbereichs Kunst und Design der Fachhochschule Köln
Beziehungen zu Organisationen	Fachhochschule Köln (-1986)
Systematik	13.4p Personen zu Malerei, Zeichnung, Grafik
Typ	Person (piz)
Thema in	2 Publikationen 1. <i>Skizzenbücher</i> <i>Marx, Karl. - Köln : Salon-Verl., 2000</i> 2. <i>Karl Marx, Gemälde</i> <i>Köln : DuMont, 1994</i>

Beispiel: <http://d-nb.info/gnd/118578545>¹

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
```

```
@prefix gnd: <http://d-nb.info/standards/elementset/gnd#>
```

```
@prefix owl: <http://www.w3.org/2002/07/owl#> .
```

```
<http://d-nb.info/gnd/118578545>
```

```
  gnd:preferredNameForThePerson "Marx, Karl" ;
```

```
  gnd:professionOrOccupation
```

```
    <http://d-nb.info/gnd/4033423-5> ;
```

```
  foaf:page
```

```
    <http://de.wikipedia.org/wiki/Karl_Marx_%28Maler%29> ;
```

```
  owl:sameAs <http://viaf.org/viaf/96119561> .
```

```
<http://d-nb.info/gnd/4033423-5>
```

```
  gnd:preferredNameForTheSubjectHeading "Künstler" .
```

9

- ▶ Alle Dinge werden mit einer URI identifiziert
- ▶ Alle Daten bestehen aus einzelnen Aussagen (**Triples**)
 - ▶ Subjekt (immer eine URI)²
 - ▶ Property (Eigenschaft identifiziert durch eine URI)
 - ▶ Objekt (Zeichenkette oder URI)²

Beispiel:

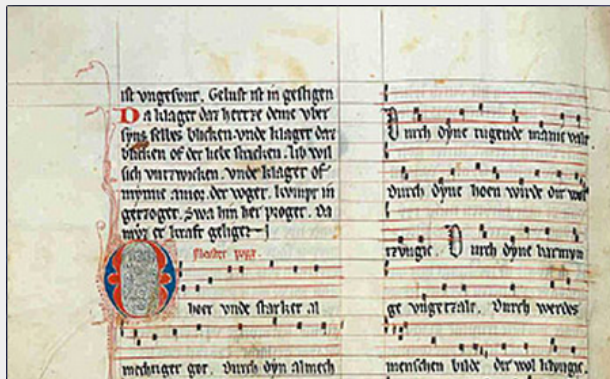
S <http://d-nb.info/gnd/11857854>
P <http://d-nb.info/standards/elementset/gnd#professionOrOccupation>
O <http://d-nb.info/gnd/4033423-5>

- ▶ Definition von Eigenschaften
- ▶ **Einheitliche** Eigenschaften
 - ▶ `gnd:preferredNameForThePerson`
 - ▶ `foaf:name` (`http://xmlns.com/foaf/0.1/name`)
 - ▶ `schema:name` (`http://schema.org/name`)
 - ▶ ...
- ▶ z.B. Katalogisierung mit RDA basiert auf einer RDF-Ontologie, andere Datenmodelle lassen sich auf Ontologien abbilden
- ▶ Ontologien sind allerdings auch mischbar!

Katalogisierung: Intellektuelle Erstellung von Daten

1. Zeichenketten oder andere Digitalisate
2. Verknüpfungen

Zeichenketten oder andere Digitalisate

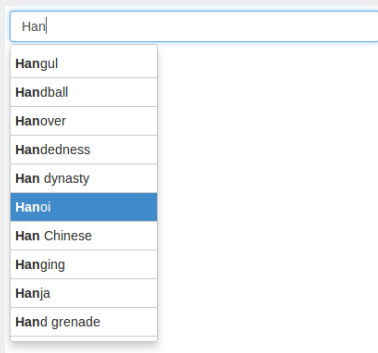


nur genau einmal notwendig, danach Verknüpfungen

13

Verknüpfungen

- ▶ Auswahl aus bereits vorhandenen URIs
- ▶ Mit geeigneten Hilfsmitteln



A screenshot of a search interface. At the top is a text input field containing the text 'Han|'. Below the input field is a dropdown menu with a list of suggestions. The suggestions are: Hangul, Handball, Hanover, Handedness, Han dynasty, Hanoi, Han Chinese, Hanging, Hanja, and Hand grenade. The 'Hanoi' option is highlighted with a blue background.

Hangul
Handball
Hanover
Handedness
Han dynasty
Hanoi
Han Chinese
Hanging
Hanja
Hand grenade

- ▶ Vorteile

- ▶ Einheitliches Datenformat
- ▶ Verfügbarkeit der Daten
- ▶ Flexibleres Zusammenführen und Ausschnitte bilden

⇒ Mehrwert durch *Zusammenarbeit*

- ▶ LOD macht keinen Sinn wenn...

- ▶ Alle Daten selber erfasst und genutzt werden sollen

⇒ Zusammenarbeit nicht gewünscht ist

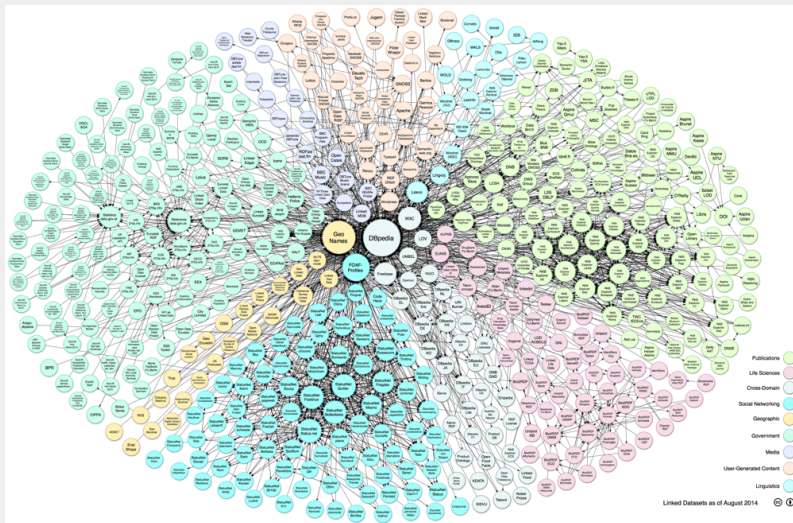
Konzeptuelle Nachteile

- ▶ Nicht alle Daten lassen sich gleich gut in RDF ausdrücken (z.B. Hierarchien, Reihenfolgen. . .)
- ▶ Datenmodellierung ist kompliziert weil die Dinge kompliziert sind

Ggf. reichen auch erstmal andere Formate, sofern URIs dabei sind (CSV-Tabellen, BEACON-Linkdumps. . .)

- ▶ Alles lässt sich mit allem verknüpfen (mittels URIs)
- ▶ Katalogisierung ist Verknüpfung oder Digitalisierung
- ▶ Tripel als kleinste Dateneinheit
 - ▶ leichte Nachnutzung und Zusammenführung
 - ▶ Gemeinsame Nutzung von RDF-Eigenschaften & Ontologien
- ▶ Massenweise Verknüpfungen

Linking Open Data cloud diagram



SWIB Semantic Web in Libraries (seit 2009)

<http://swib.org/>

LODLAM Linked Open Data in Libraries, Archives, and Museums (seit 2011)

<http://lodlam.net>

Linking Open Data cloud diagram (2014) by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak <http://lod-cloud.net/>

