Connections that work
Linked Open Data demystified

Jakob Voß

October 22th, 2014 (#kivepa2014)
connections
everything is connected
everything is connected with everything
six degrees of separation
Linked Open Data cloud
What is a useful connection?
What is a useful connection?
connections that work
connection that can be processed
connection that can be processed automatically
connection expressed as Linked Open Data
Overview

1. History
2. Present
3. Future
History
Henriette Avram (1919-2006)
<table>
<thead>
<tr>
<th>Description</th>
<th>Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Entry</td>
<td>100</td>
</tr>
<tr>
<td>Filing Title</td>
<td>150</td>
</tr>
</tbody>
</table>

### Statements

<table>
<thead>
<tr>
<th>Title</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edition</td>
<td>250</td>
</tr>
<tr>
<td>Imprint</td>
<td>300</td>
</tr>
<tr>
<td>Collation</td>
<td>400</td>
</tr>
</tbody>
</table>

### Notes

<table>
<thead>
<tr>
<th>Series-Add</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series-No</td>
<td>510</td>
</tr>
<tr>
<td>Notes</td>
<td>600</td>
</tr>
</tbody>
</table>

### Tracings

<table>
<thead>
<tr>
<th>Subject</th>
<th>700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pers Auth</td>
<td>710</td>
</tr>
</tbody>
</table>

#### Corp Auth:

- Govt Body: 72B
- Soc or Inst: 72C
- Relig Body: 72D
- Miscell: 72E

### Uniform

<table>
<thead>
<tr>
<th>Title</th>
<th>740</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>750</td>
</tr>
<tr>
<td>Copy Stat</td>
<td>800</td>
</tr>
<tr>
<td>Nat Bib No</td>
<td>830</td>
</tr>
<tr>
<td>NBN (over 15)</td>
<td>831</td>
</tr>
<tr>
<td>LC Call No</td>
<td>900</td>
</tr>
<tr>
<td>DDC No</td>
<td>920</td>
</tr>
<tr>
<td>LC Card No</td>
<td>940</td>
</tr>
</tbody>
</table>

### Library of Congress

- Call No: 533.2
- LC Card No: 940

---

The MARC pilot project (1968)
Machine Readable Cataloging (MARC)

good...

- encoding of bibliographic information as data
- controlled collaboration
Machine Readable Cataloging (MARC)

good...

- encoding of bibliographic information as data
- controlled collaboration

...but...

- primary use case: printed cards
Information Management: A Proposal (1989)
World Wide Web (WWW)

good...

- worldwide connections
- uncontrolled collaboration
World Wide Web (WWW)

good...

- worldwide connections
- uncontrolled collaboration

...but...

- only for browsing interfaces
Present
Linked Open Data cloud
Semantic Web

good...

- data integration and analysis
- semi-controlled collaboration
Semantic Web

good...

▶ data integration and analysis
▶ semi-controlled collaboration

...but...

▶ promises of artificial intelligence
Linked Open Data (LOD)

good…

▶ data integration and analysis
▶ semi-controlled collaboration

…but…

▶ promises of artificial intelligence
Linked Open Data (LOD)

sounds nice...

- encoding of information as data
- worldwide connections
- data integration and analysis
Linked Open Data (LOD)

sounds nice...

▶ encoding of information as data
▶ worldwide connections
▶ data integration and analysis

...but...

▶ how does it actually work?!
a little theory of data
Data

- all data refers to something
Data

- all data refers to something
- some data refer to the “same” thing
Data

- all data refers to something
- some data refer to the “same” thing
- for instance the same person
  - as author in a library catalog
  - in a university research management system
  - Wikipedia article about the person
  - ...
Data

- all data refers to something
- some data refer to the “same” thing
- Linked Data makes true sense only with data from multiple sources about the same thing
“Old School” Library Linked Data

authority files registry, classification, thesaurus...
controlled vocabularies no homonyms/synonyms
identifier notations, codes, numbers...
“Old School” Library Linked Data

authority files registry, classification, thesaurus...
controlled vocabularies no homonyms/synonyms
identifier notations, codes, numbers...
same identifier \(\Rightarrow\) data refers to the same thing
unique identification
unique identification with URIs
an example
Karl Marx
Karl Marx (German artist, 1929-2008)
unique identification of things

- “author:”
  [http://d-nb.info/gnd/118578545](http://d-nb.info/gnd/118578545)
unique identification of things

- **object:**
  
  http://d-nb.info/gnd/118578545
  ("Karl Marx, the artists")
unique identification of things and connections

- **object:**
  http://d-nb.info/gnd/118578545 ("Karl Marx, the artists")

- **property:**
  http://purl.org/dc/terms/creator ("authorship, as defined by Dublin Core")
Ei kaksi kolmennettta
Ei kaksi kolmannetta

- **subject:**
  http://d-nb.info/940697734 ("a specific book with paintings")

- **property:**
  http://purl.org/dc/terms/creator ("authorship, as defined by Dublin Core")

- **object:**
  http://d-nb.info/gnd/118578545 ("Karl Marx, the artists")
Eiaksi kolmannettta

- **subject:**
  http://d-nb.info/940697734
  ("a specific book with paintings")

- **property:**
  http://purl.org/dc/terms/creator
  ("authorship, as defined by Dublin Core")

- **object:**
  http://d-nb.info/gnd/118578545
  ("Karl Marx, the artists")

**hey, that’s an RDF triple!**
technical break
An RDF triple (N-Triples syntax)

<http://d-nb.info/940697734>
<http://purl.org/dc/terms/creator>
@prefix dct: <http://purl.org/dc/terms/> .

An RDF triple (aREF syntax)

http://d-nb.info/940697734:
dc_creator:
    http://d-nb.info/gnd/118578545
An RDF triple (aREF syntax)

http://d-nb.info/940697734: 
dc_creator:
    http://d-nb.info/gnd/118578545

RDF syntax converters exist!
An RDF triple (aREF syntax)

http://d-nb.info/940697734:
  dc_creator:
    http://d-nb.info/gnd/118578545

RDF syntax converters exist!
always use the RDF syntax you understand best!
An RDF triple (aREF syntax)

http://d-nb.info/940697734:
dc_creator:
   http://d-nb.info/gnd/118578545

RDF syntax converters exist!
always use the RDF syntax you understand best!
(nobody understands RDF/XML syntax)
Multiple RDF triples (Turtle syntax)

@prefix dct: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

<http://d-nb.info/940697734>
dct:title "Gemälde" ;

<http://d-nb.info/gnd/118578545>
foaf:name "Karl Marx" .
That’s RDF!

- all things identified by URIs\(^1\)

\(^1\)or a blank node
That’s RDF!

- all things identified by URIs$^1$
- all data consists of triples
  - subject (always an URI)$^1$
  - property (always an URI)
  - object (URI or character string)$^1$

$^1$or a blank node
Ontologies

- set of common properties and rules
- foaf:name ([http://xmlns.com/foaf/0.1/name](http://xmlns.com/foaf/0.1/name))

More specialized ontologies exist:
- schema.org
- RDA-ontology and BIBFRAME

Ontologies can be mapped and merged.
Ontologies

set of common properties and rules

- foaf:name
  (http://xmlns.com/foaf/0.1/name)
- dct:title
  (http://purl.org/dc/terms/title)
Ontologies

set of common properties and rules

- foaf:name
  (http://xmlns.com/foaf/0.1/name)
- dct:title
  (http://purl.org/dc/terms/title)

more specialized ontologies exist

- schema.org
- RDA-ontology and BIBFRAME
- ...

ontologies can be mapped and merged
Ontologies

set of common properties and rules

- foaf:name
  (http://xmlns.com/foaf/0.1/name)
- dct:title
  (http://purl.org/dc/terms/title)

more specialized ontologies exist

- schema.org
- RDA-ontology and BIBFRAME
- ...

ontologies can be mapped and merged
end of technical break.

wake up!
Lesson learned

- RDF data consists of triples
Lesson learned

- RDF data consists of triples
- subjects, properties, and most objects identified by URIs
Lesson learned

- RDF data consists of triples
- subjects, properties, and most objects identified by URIs
- common properties from ontologies
Lesson learned

- RDF data consists of triples

- subjects, properties, and most objects identified by URIs

- common properties from ontologies

→ easy to mix and merge!
summary of the present linked open data
Linked Open Data

1. **Data**
   with URIs, in RDF
Linked Open Data

1. **Data**
   with URIs, in RDF

2. **Open**
   accessible via HTTP-URIs
Linked Open Data

1. **Data**
   with URIs, in RDF

2. **Open**
   accessible via HTTP-URIs

3. **Linked**
   contains other HTTP-URIs
LINKED DATA
- On the web
- Open license
- Machine-readable data
- Non-proprietary format
- RDF standards
- Linked RDF

5 stars: Is your data 5 stars?

Kippis!
Linked RDF
Linked RDF contains other HTTP-URIs.
Future
Future of cataloging
cataloging?
intellectual creation of data
intellectual creation of data about existing works
Creation of data about existing works

1. digitization
Digitization
required only once per work
creation of data about existing works

1. digitization
creation of data about existing works

1. digitization

2. connections
creation of data about existing works

1. digitization

2. connections

infinite combination of connections
Connections that work

- selection from already existing URIs
Connections that work

- selection from already existing URIs
- with appropriate tools
why Linked Open Data?
Benefits

- common data format
- accessibility of data
- flexible aggregation and subsets
- collaborative creation
Disadvantages

- some data is more difficult to express in RDF (e.g. hierarchies and order)
Disadvantages

▶ some data is more difficult to express in RDF (e.g. hierarchies and order)
▶ data modeling is complicated
Disadvantages

- some data is more difficult to express in RDF (e.g. hierarchies and order)
- data modeling is complicated
- reality is complicated
Disadvantages

- some data is more difficult to express in RDF (e.g. hierarchies and order)
- data modeling is complicated
- reality is complicated
- collaboration can be complicated
where to start?
Linked Data Finland

http://www.ldf.fi/

please contact to share your datasets!
summary
everything can be connected
everything can be connected with URIs
everything can be connected as soon as you connect it!
Conferences

**SWIB**  Semantic Web in Libraries (since 2009)
http://swib.org/

**LODLAM**  Linked Open Data in Libraries, Archives, and Museums (since 2011)
http://lodlam.net

**SEMANTiCS**  http://www.semantics.cc
Image sources

- Six degrees of separation by Daniel Walker (User:Dannie-walker)
  http://lod-cloud.net/
- Henriette Avram from
  http://www.loc.gov/loc/lcib/0605/avram.html
  http://www.w3.org/History/1989/proposal.html
- Tim Berners-Lee (2005) by Uldis Bojārs from Flickr
thanks!
questions!