

# Open Data / Linked Open Data

**Herbsttagung der BibliothekarInnen der MPG  
03.11.2011**

## 1. Definitionen

- Open Data
- Linked Open Data
- Semantic Web

## 2. Beispiele

**Motivation / Warum?**

## Open (Bibliographic) Data

- **Januar 2010:** The CERN Library publishes its book catalog as Open Data
- **März 2010:** Freigabe der Katalogdaten
  - Vorreiter Universität- und Stadtbibliothek Köln
  - und weitere Bibliotheken in NRW und Rheinland-Pfalz in Zusammenarbeit mit dem HBZ (im internen Format)
- **Mai 2010 folgen**
  - Universitätsbibliotheken Konstanz, Tübingen und Mannheim in Zusammenarbeit mit dem BSZ (im MARC21-Format)



## Motivation?



**Der freie Zugang zu wissenschaftlicher Information**

<http://open-access.net/>



<http://creativecommons.org/>



<http://de.creativecommons.org/>

## Creative Commons (CC):

- Non-Profit-Organisation
- Bietet vorgefertigter Lizenzverträge für die Veröffentlichung und Verbreitung digitaler Medieninhalte (Hilfestellung)

-  Namensnennung ([Details](#))
-   Namensnennung-KeineBearbeitung ([Details](#))
-   Namensnennung-NichtKommerziell ([Details](#))
-    Namensnennung-NichtKommerziell-KeineBearbeitung ([Details](#))
-    Namensnennung-NichtKommerziell-Weitergabe unter gleichen Bedingungen ([Details](#))
-   Namensnennung-Weitergabe unter gleichen Bedingungen ([Details](#))

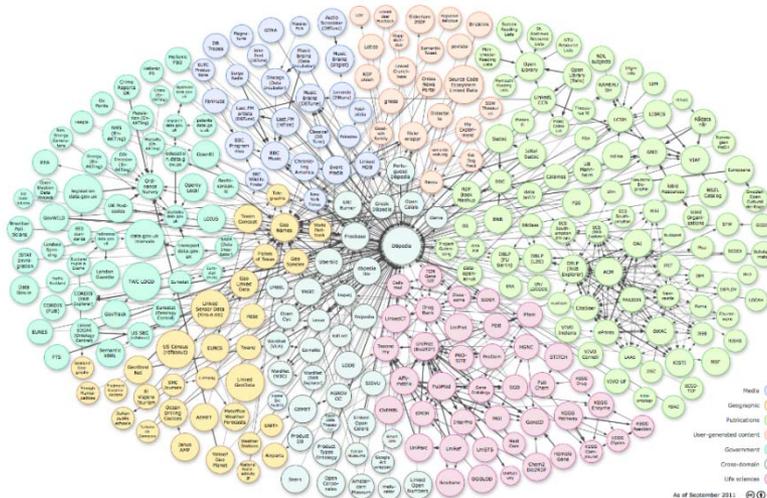
<http://de.creativecommons.org/was-ist-cc/>

<http://creativecommons.org/licenses/>

## Linked Open Data (LOD)

Mechanismus, um Daten im Web miteinander in Beziehung zu setzen

- bezeichnet im WWW frei verfügbare Daten,
- die per URI identifiziert sind und
- darüber direkt per HTTP abgerufen werden können und
- ebenfalls per URI auf andere Daten verweisen.



Linked Open Data Cloud

erstmalig: Oktober 2007

Stand September 2011

<http://richard.cyganiak.de/2007/10/lod/>



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>

## 1. Definition:

- *The term “Semantic Web” refers to W3C’s vision of the Web of linked data.*
- *Semantic Web technologies enable people to create data stores on the Web,*
- *build vocabularies, and*
- *write rules for handling data.*
- *Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS.*



## 2. Definition:

*„Das Semantic Web ist eine Vision von der Erweiterung des Internets, um die im WWW verfügbaren Informationen mit für Maschinen bearbeitbarer Semantik zu erweitern.“*



*Es handelt sich dabei um eine offene, aktive und v.a. vom W3C vorangetriebene Forschungsunternehmung, die im Zusammenhang mit der Diskussion um Metadaten oder Schlagworte, wie. z.B. RDF, XML, etc. gesehen werden muss.“*

## 1. HTML:

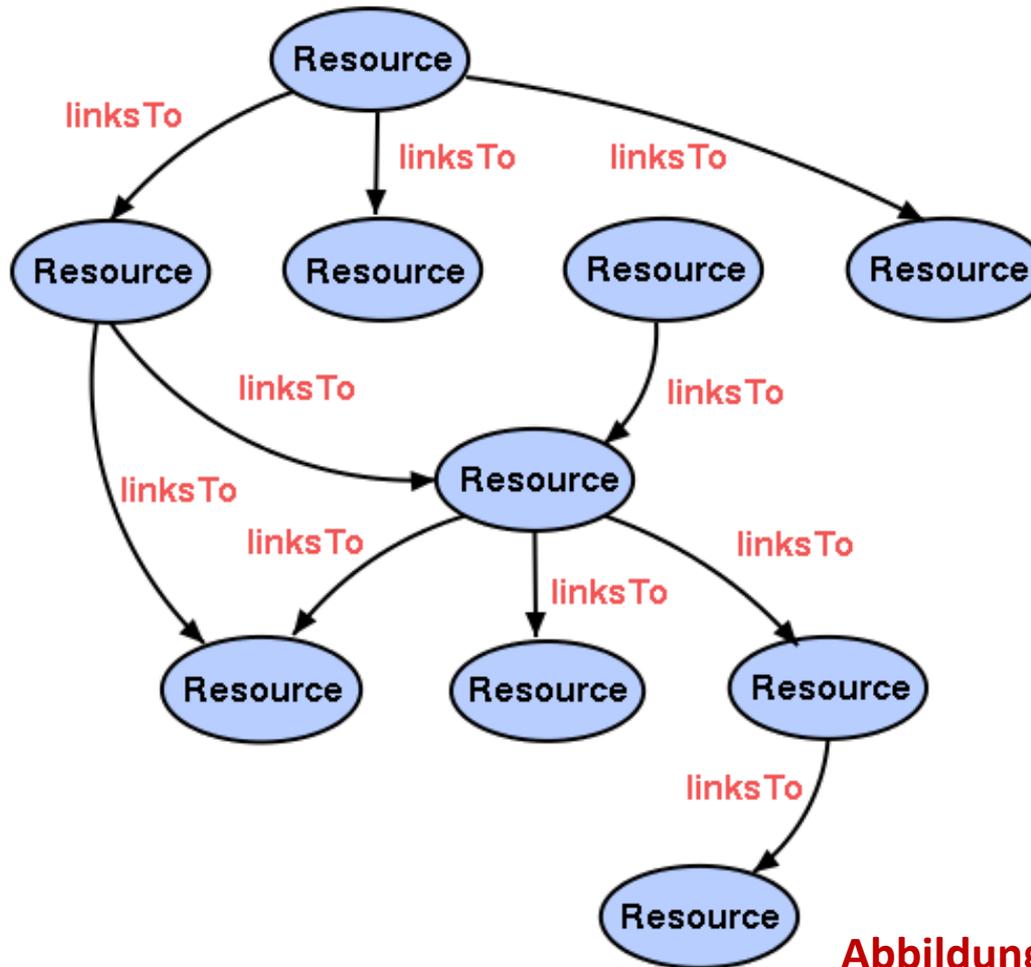
als Dokumentenbeschreibungssprache, die festlegt, wie die Information gegliedert ist und wie die Dokumente verknüpft sind (Hyperlinks).

## 2. HTTP:

als Protokoll, mit dem der Browser Informationen vom Webserver anfordern kann.

## 3. URL:

als eindeutige Adresse bzw. Bezeichnung einer Ressource (z. B. einer Webseite), die in Hyperlinks verwendet wird.



**Abbildung: World Wide Web aus W3C**

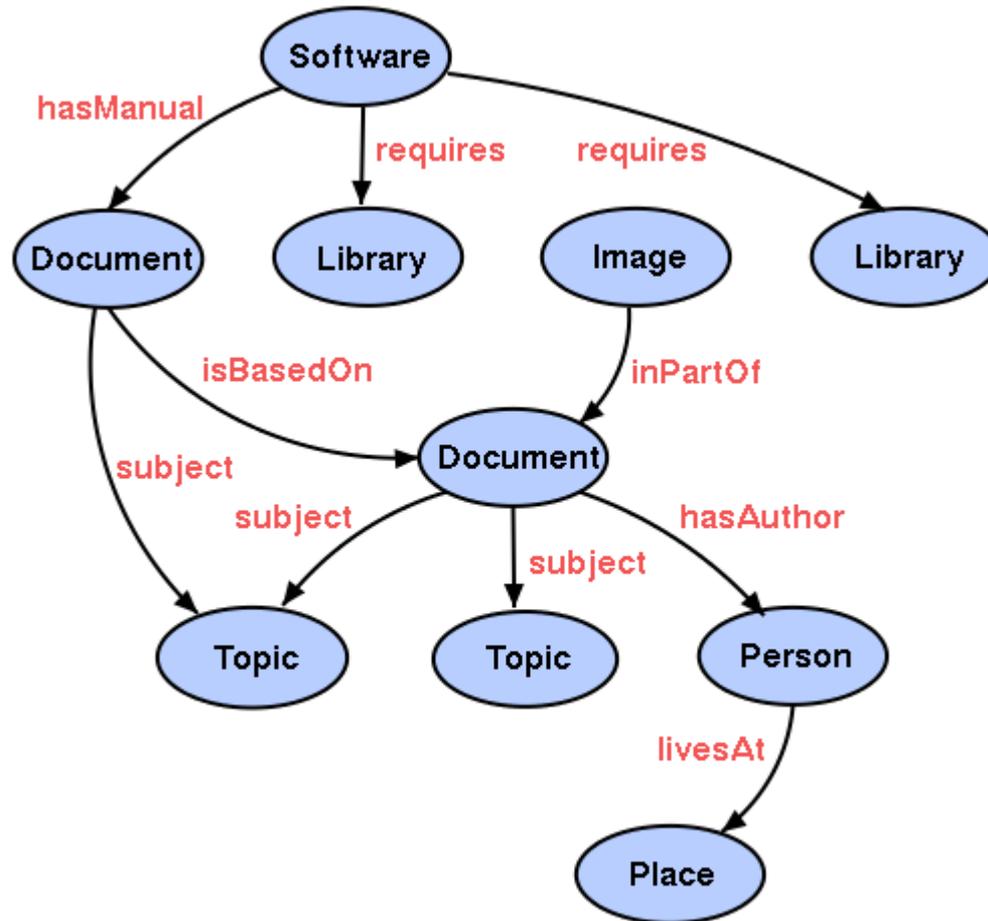
- **Repräsentation auf den Menschen als Enduser ausgerichtet.**
- **Menschliche Nutzer kann Informationen in Beziehung setzen**
- **Typisches, sehr einfaches Beispiel:**

Suche nach Kohl:  
Gemüse? Altkanzler? etc.

- **und: Informationsflut wächst und wächst**

- **die Bedeutung von Informationen für Computer verwertbar machen,**
- **d.h. die Informationen im Web sollen von Maschinen interpretiert und automatisch maschinell weiterverarbeitet werden können.**
  - Informationen über Orte, Personen und Dinge sollen mit Hilfe des Semantischen Webs von Computern miteinander in Beziehung gesetzt werden können.
  - Ermöglichen neuartige Anwendungen und Dienste
  - Daten können für vorher nicht notwendigerweise vorgesehene Zwecke wieder- und weiterverwendet werden
  - **1. Idee: Einsatz von künstlicher Intelligenz**  
(derzeit: schlechte Ergebnisse von zuverlässigen Anwendungen)

- **Einheitliche, offene Standards für die Beschreibung der Informationen**
  - Informationen zwischen Anwendungen und Plattformen austauschen und in Beziehung setzen können (Interoperabilität)
  - Neben formal klarer Definition aber auch Flexibilität und Erweiterbarkeit notwendig.
- **Standards sind (Informations-Spezifikationssprachen):**
  - (XML) (syntaktische Grundlage)
  - RDF(S) (S für Schema)
  - OWL
  - RDF(S) und OWL sind Ontologiesprachen  
(Ontologie <-> Wissensbasis)
- **Formulierung von Regeln, mit deren Hilfe das Computersystem Rückschlüsse bilden kann**



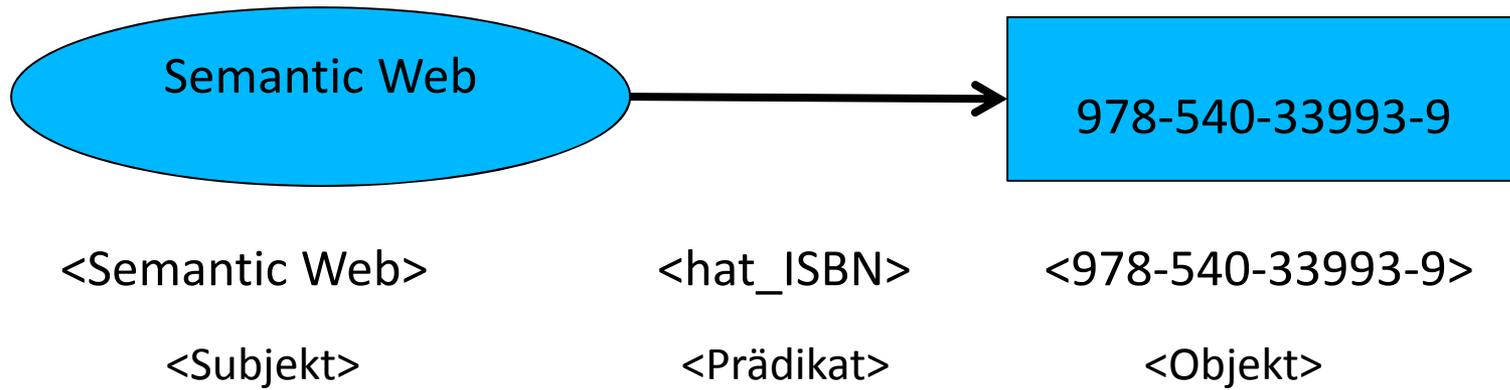
**Abbildung: Semantic Web aus W3C**

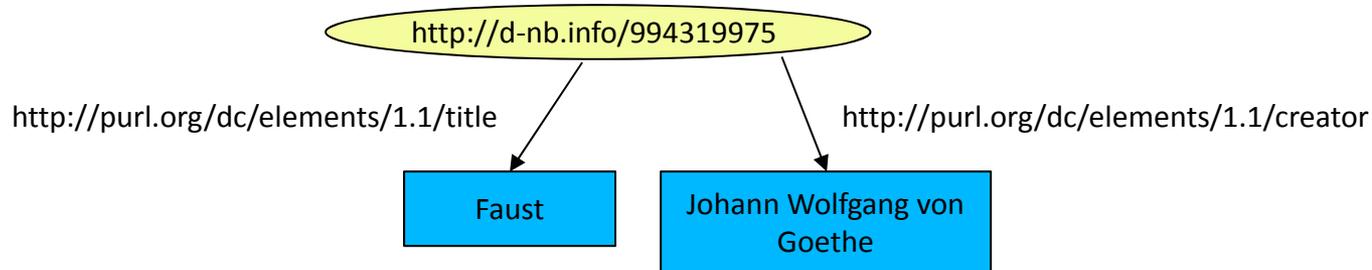
- **Beschreibungssprache für Informationen einer Webressource**
- **Auszeichnungssprache für Metadaten**
- **Basiert auf sogenannten Tripeln**
  - **Subjekt**
  - **Prädikat (Eigenschaft)**
  - **Objekt (Eigenschaftswert)**

## Beispiele:

<Semantic Web> <hat\_ISBN> <978-540-33993-9>  
<Subjekt> <Prädikat> <Objekt>

<Semantic Web> <hat\_ISBN> <978-540-33993-9> ∩  
<hat\_Autor> <Hitzler, Pascal>





```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
```

```
  <rdf:Description rdf:about="http://d-nb.info/994319975 ">
```

```
    <dc:title>Faust</dc:title>
```

```
    <dc:creator>Johann Wolfgang von Goethe</dc:creator>
```

```
  </rdf:Description>
```

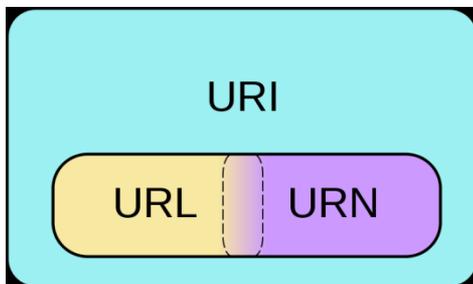
```
</rdf:RDF>
```

[http://www.d-nb.de/hilfe/service/linked\\_data\\_service.htm](http://www.d-nb.de/hilfe/service/linked_data_service.htm)

## Eindeutiger Bezeichner

### Beispiele:

- <http://www.bsz-bw.de>
- <mailto:volker.conradt@bsz-bw.de>
- <http://swb.bsz-bw.de/DB=2.1/PPNSET?PPN=256268142> 
- urn:nbn:de:bsz:boa-bsz3134565691
- [http://opus.bsz-bw.de/swop/volltexte/2009/771/pdf/LZA\\_Schichtenmodell\\_BSZ\\_Koll2009\\_Wolf.pdf](http://opus.bsz-bw.de/swop/volltexte/2009/771/pdf/LZA_Schichtenmodell_BSZ_Koll2009_Wolf.pdf)



### Cool URIs don't change

- What makes a cool URI?
- A cool URI is one which does not change.
- What sorts of URI change?
- *URIs don't change: people change them.*

Tim Berners Lee

Tom Gruber:

**RDF:**

**legt Syntax fest**

**Ontologie: Vokabular mit Regeln**

*A specification of a representational vocabulary for a shared domain of discourse — definitions of classes, relations, functions, and other objects — is called an **ontology**.*

An **ontology** is an explicit specification of a conceptualization.

- **Darstellung komplexer Wissensbeziehungen**
- **Netz von Hierarchien, in dem Informationen über logische Beziehungen miteinander verknüpft sind oder sein könnten.**
- **Ontologie <-> Wissensbasis**
- **OWL: Web Ontology Language**  
**Beschreibungssprache für Klassen und Relationen**  
**OWL ist die zurzeit populärste Sprache für die Modellierung von Ontologien**



<http://www.w3.org/2005/Incubator/lld/wiki/Vocabularies>

Property	Range	Value String
dcterms:title	literal	YES
dcterms:created	literal	YES
dcterms:language	non-literal	YES
dcterms:subject	non-literal	YES
dcterms:creator	non-literal	YES
foaf:firstName	literal	YES
foaf:family_name	literal	YES
foaf:mbox	non-literal	no

foaf: friend of a friend

## Buchhändler:

<http://www.springer.com/computer/database+management+&+information+retrieval/book/978-3-540-33993-9>

onix:productNumber "978-3-540-33993-9"

onix:productTitle "Sematic Web"

onix:author "Pascal Hitzler"

onix:publisher "Springer Verlag"

## KatalogisiererInnen

<http://www.worldcat.org/oclc/188085465>

<http://spica.bsz-bw.de:5580/DB=2.1/PPNSET?PPN=310658349/>

dc:title "Semantic Web"

dc:creator "http://www.viaf.org/viaf/55340633"

dc:publisher "Berlin ; Heidelberg : Springer, 2008"

frbr:manifestation "http://spica.bsz-bw.de:5580/DB=2.1/ PPNSET?PPN=276343212/"

## Bibliothek

frbr:item "http://www3.biss.belwue.de/pica\_opac/plr\_biss.php?wcmu=pollux&HLIB=593&ppn=256268142&epn=555966402"

## Buchhändler:

<http://www.springer.com/computer/database+management+&+information+retrieval/book/978-3-540-33993-9>

onix:productNumber "978-3-540-33993-9"

onix:productTitle "Sematic Web"

onix:author "Pascal Hitzler"

onix:publisher "Springer Verlag"

## KatalogisiererInnen

<http://www.worldcat.org/oclc/188085465> >

<http://spica.bsz-bw.de:5580/DB=2.1/PPNSET?PPN=310658349>

dc:title "Semantic Web"

dc:creator "http://www.viaf.org/viaf/55340633"

dc:publisher "Berlin ; Heidelberg : Springer, 2008,,

frbr:manifestation "http://spica.bsz-bw.de:5580/DB=2.1/ PPNSET?PPN=276343212/"

## Bibliothek

frbr:item "http://www3.biss.belwue.de/pica\_opac/plr\_biss.php?wcmu=pollux&HLIB=593&ppn=256268142&epn=555966402"

## Nutzer

rdfs:seeAlso "http://de.wikipedia.org/wiki/Semantisches\_Web "



**Anfrage:**

**Finde die Namen aller afrikanischen Hauptstädte  
und das Land, in dem sich jedes befindet.**

PREFIX abc: <http://example.com/exampleOntology#>

```
SELECT ?capital ?country
```

```
WHERE
```

```
{
```

```
?x abc:cityname ?capital.
```

```
?y abc:countryname ?country.
```

```
?x abc:isCapitalOf ?y.
```

```
?y abc:isInContinent abc:africa.
```

```
}
```

## 1. Definitionen

- Open Data
- Linked Open Data
- Semantic Web

## 2. Beispiele

**Motivation / Warum?**



CC0 1.0

<http://www.bl.uk/bibliographic/datafree.html#>

## Linked Open BNB in RDF/XML:

[Bibliographic Ontology](#)

[Bio: A Vocabulary for Biographical Information](#)

[British Library Terms](#)

[Dublin Core](#)

[Event Ontology](#)

[FOAF: Friend of a Friend](#)

[Org: An Organization Ontology](#)

[OWL](#)

[SKOS](#)

[RDF Schema](#)

[WGS84 Geo Positioning](#)

and the preliminary versions of the [ISBD](#) and [RDA](#)

## BNB in Basic RDF/XML:

[Dublin Core](#)

[SKOS](#)

[OWL](#), Time Ontology in OWL

[RDF Schema](#)

[Bibliographic Ontology](#)

and the preliminary versions of the [ISBD](#) and

[RDA](#)

- data model [version 1.01, 1 August 2011]
- draft schema [version 1.01, 28 July 2011]

## Dokumentation des Linked Data Services der DNB

<https://wiki.d-nb.de/display/LDS/Dokumentation+des+Linked+Data+Services+der+DNB>

- Beispiel (Ausschnitt aus der Datenbasis, ca. 500 KB, Stand 04/2010): <http://tinyurl.com/39zy2tw>
  - Personen (PND), Stand 04/2011: [Teil 1](#) (RDF/XML, ca. 265 MB ZIP File), [Teil 2](#) (RDF/XML, ca. 5 MB ZIP File)
  - Körperschaften (GKD), Stand 04/2011: [Teil 1](#) (RDF/XML, ca. 95 MB ZIP File), [Teil 2](#) (RDF/XML, ca. 16 MB ZIP File)
  - Sachschlagworte (SWD), Stand 07/2011: [Alle Daten](#) (RDF/XML, ca. 14 MB ZIP File)
- 
- [Linked Data Service - Dokumentation](#)
  - [Nutzungsbedingungen der Datendienste der Deutschen Nationalbibliothek](#)
  - [Linked Data Service - FAQ](#)

The logo for culturegraph, featuring the word "culturegraph" in white lowercase letters on a blue rectangular background.

<http://www.culturegraph.org/>

- **Projekt: Aufbau von Linked-Open-Data-Services**
- **Projektbeteiligte: DNB, HBZ und AG-Verbundsysteme**
  
- **Aufbau von gemeinsamen Identifikatoren (Uniform Resource Identifier) für kulturelle Erzeugnisse (Bücher und andere Texte, Gemälde, Skulpturen, Musikstücke usw.)**



<http://www.europeana.eu/portal/>

**Metadaten:** CCO

**The Europeana Data Model for Cultural Heritage**

### **Terms of Use:**

“The website "Europeana.eu" is run by the Europeana Foundation and gives access to a wide range of digitised cultural heritage across Europe.

This material is provided to the Europeana Foundation by a large number of contributing institutions and organisations, which in many cases, own rights in the information provided.”

## Deutsche Digitale Bibliothek

### Portal für Kultur und Wissenschaft

<http://www.deutsche-digitale-bibliothek.de/>

#### Sicherung des Urheberrechts

Die Deutsche Digitale Bibliothek wahrt selbstverständlich die Urheber- und Verwertungsrechte. Zum Start werden zunächst ausschließlich urheberrechtsfreie Werke eingestellt. Geplant ist, dass künftig Urheber und Verlage für die Internet-Veröffentlichung ihrer Werke über die Deutsche Digitale Bibliothek einen angemessenen Preis verhandeln werden können. Wie dieses Verfahren genau ablaufen soll, wird auch in Kooperation mit dem Börsenverein des Deutschen Buchhandels und der Verwertungsgesellschaft Wort ausgearbeitet.

#### In Diskussion:

- **Metadaten (CC0)**
- **Erschließungsdaten (nicht CC0)**

„Der **Verbundkatalog B3Kat** des BVB und KOBV mit fast **22 Millionen Titeldaten** wird im **Spätherbst 2011** sowohl **Open Data** (= uneingeschränkte Bereitstellung im Internet) als auch **Linked Open Data** (= integrierbar in das offene Semantic Web) verfügbar gemacht werden.“

Sept. 2011 / KOBV-Forum Berlin  
Eginhardt Scholz, BSB/BVB

[http://www.kobv.de/ueber den kobv/kobv forum/9 kobv forum 2011/](http://www.kobv.de/ueber_den_kobv/kobv_forum/9_kobv_forum_2011/)

## **Empfehlungen des KOBV-Kuratoriums zur Katalogdatenfreigabe an die KOBV-Mitgliedsbibliotheken**

verabschiedet: 20.05.2011

- Bibliographische Metadaten können einen wesentlichen Beitrag für das sog. Semantic Web liefern.
- Open Bibliographic Data erhöhen wesentlich die Sichtbarkeit von Bibliotheken und Bibliotheksbeständen.
- Linked Data haben das Potential für innovative und bessere Informationsangebote.
- Die bisher entstandenen Kosten, die in die Kataloge eingeflossen sind, sind hoch, aber der Wert der Katalogdaten erhöht sich durch Linked Data.
- Die bilaterale Abgabe von Katalogdaten an Dritte ist bisher schon üblich; mit Open Data würde diese Form der Kooperation (mit geringerem Aufwand) konsequent fortgeführt.

- In einem ersten Schritt können Datensätze und -elemente, für die eine endgültige rechtliche Klärung noch nicht vorliegt, ausgeklammert werden. Teilnehmerbibliotheken sollten die Möglichkeit haben, ihre Bestände von der Freistellung auszunehmen.
- Da eine optimale Nutzung nur durch eine uneingeschränkte Freigabe erzielt werden kann, wird die Verwendung der Creative Commons Lizenz CC0 empfohlen.
- Empfohlen wird des weiteren, die Daten in Form von RDF-Tripeln und zusätzlich in einem Bibliotheksformat (vorzugsweise MARCXML), alternativ in MODS, bereitzustellen, damit die Nutzung im Sinne von Linked Data möglichst effektiv ist. Die Tripel-Bildung auf der Basis der verschiedensten Ontologien und Vokabularien (Dublin Core, Biblioontology, MARC Ontology, MODS Ontology, RDA-Vokabular, FRBR Ontology usw.) beschränkt sich dabei in der Regel auf ein Kernset von Datenelementen.
- Eine einmalige Bereitstellung ist nicht ausreichend. Um den größtmöglichen Nutzen zu erzielen, müssen die bereitgestellten Daten regelmäßig aktualisiert werden (Korrekturen, Neuaufnahmen).

# Open Data / Linked Open Data

Herbsttagung der BibliothekarInnen der MPG  
03.11.2011

**SWIB11**  
SEMANTIC WEB  
IN BIBLIOTHEKEN  
28.-30.11. in Hamburg

**Vielen Dank für Ihre Aufmerksamkeit!**

**Fragen**

