



hbz

Wissen. Information. Innovation.

Publishing Aleph data as Linked Open Data

IGeLU , Haifa 12.9.2011

Dr. Silke Schomburg

Agenda

1. Background
2. Four Steps to Publishing Linked Open Data
3. What's Ex Libris supposed to do?

1. Background

Motivation

- Use of international and cross-domain **web standards** leads to
 - Increased **discoverability**
 - **Multiple usability** of one dataset
 - Maximized **interoperability & re-usability**
 - **Flexible** data model

Expected ROI

- **Synergy effects** through standardization of technology stack:
 - **Quantitative**: saved resources
 - **Qualitative**: improved services

A **sustainable** and consequent strategy of **publishing and linking data** on the web requires the data to be **open***.

* <http://www.opendefinition.org/>




2. Four Steps to Publishing Linked Open Data

Step 1: Find cooperating libraries

- hbz provides technology for union catalog
- Data is created cooperatively
- hbz didn't publish all data on its own because of legal uncertainties
- Four union libraries joined the initiative

Step 2: Export Data

- Using Aleph Publishing module
 - Concept: bibl. records are enriched with authority & holdings Data and written to Oracle table Z00P (XML structure)
- Problems:
 - Performance issues with incremental updates
 - disproportion between incoming/outcoming data updates
 - workaround: irregular updates as full dumps

A photograph of a TED talk stage. A large blue screen at the back of the stage displays the text "Raw Data Now!". To the left of the screen, the red "TED" logo is visible. A speaker stands on the stage in the center, facing the audience. The audience is seated in the foreground, and the stage is lit with blue light.

Raw
Data
Now!

TED

Step 3: Publishing raw data

- Load to **HTTP server**
- **Describe** published datasets
- Attach an **open license**
(for recommended open data licenses see <http://www.opendefinition.org/licenses/#Data>)
- **Register** datasets at CKAN.net
- Problem with serving consistent, standardized MAB-coded data

hbz Recently published Open Data exports

 Bearbeiten  Hinzufügen ▾  Tools ▾

Hinzugefügt von [Adrian Pohl](#), zuletzt bearbeitet von [Pascal Christoph](#) am Sep 06, 2011 ([Änderung anzeigen](#)) [Kommentar einblenden](#)

[Zur Deutschen Version.](#)


On this wiki page you'll find up-to-date information about available Open Data exports from the hbz union catalog. If you have any questions please contact [semweb\[at\]hbz-nrw \[dot\]de](mailto:semweb[at]hbz-nrw[dot]de) or post a comment on this page. (You'll have to [login/register](#) first.)
 Since April 2011 there is a [changelog](#). Informations about older data exports from March 2010 can be found [here](#).
 There is also a [download statistic site](#).



All data which is available for download here is published under a [Creative Commons CC0 licence](#). The data thus is in the Public Domain which means it belongs to everybody and can be used without restrictions for any purpose.

Recent Open Data publications

Linked-Open-hbz-Data

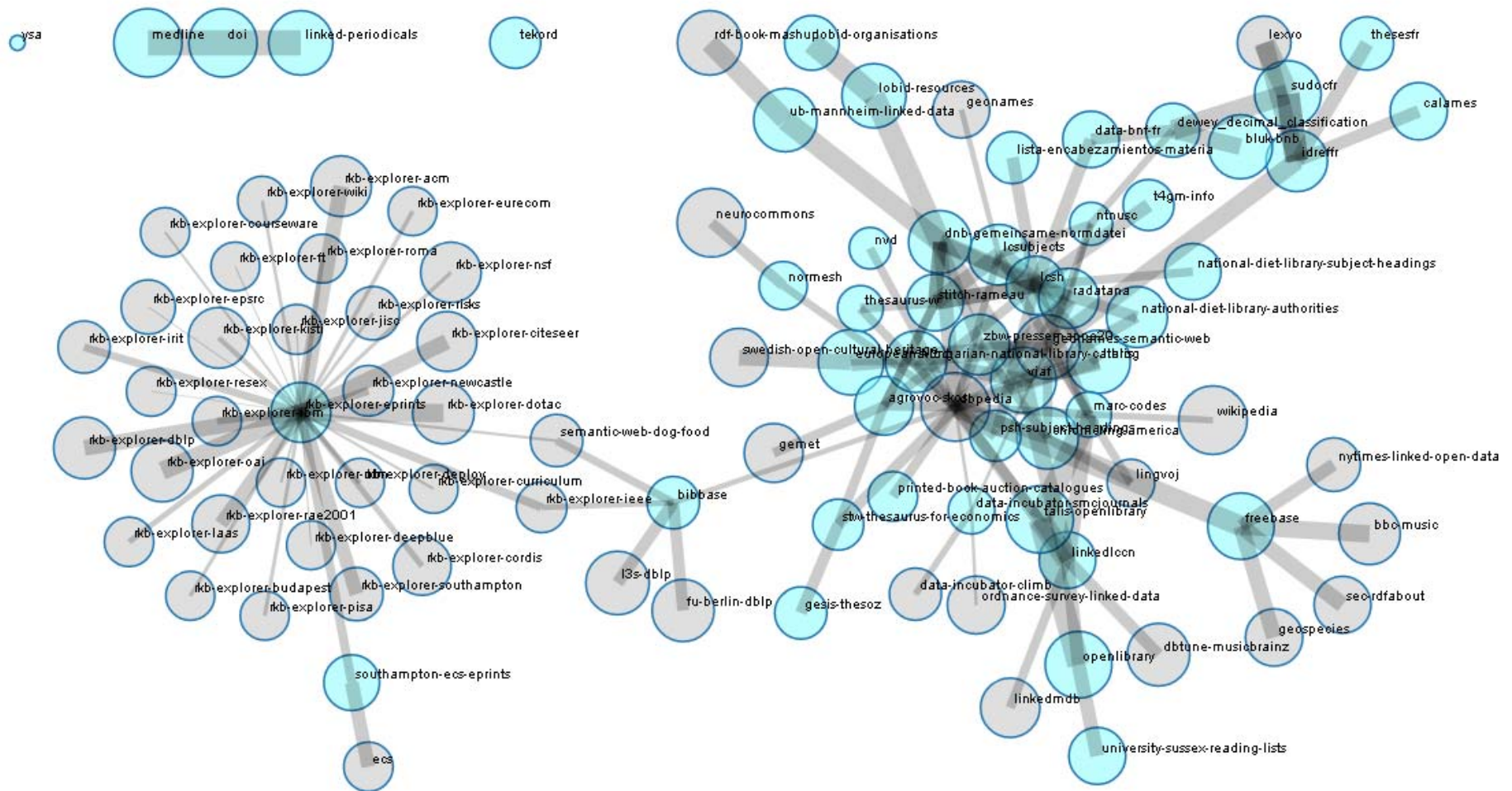
Name	Open hbz data as Linked Data
Publisher	hbz
Licence	
Export Date	2011-09-01
Publication Date	2011-09-06
Format	text/NTriples
Format documentation	See here
Extent	544 MB / 159 M triples
Checksum	f19b74eaf5cf39d556253ef0793bf017
Notes	This archive contains RDF triples of all records which have been published as Open Data.
Download URL	http://lobid.org/download/dumps/20110901_hbzlod.ttl.tar.bz2

Duisburg-Essen University Library Catalog, Campus Essen

Name	hbz records with Dortmund University Library holdings
------	---

Step 4: Identifiers & data conversion

- **Choose** suitable RDF **vocabulary** for representing bibliographic data: Bibo
- **Map** and **convert raw data** to RDF
- **Mint HTTP URIs** for described resources: <http://lobid.org/resource/{internalID}>
- **Publish RDF**
 - Triple store (4store) with SPARQL interface
 - Pubby framework for HTML representation



<http://semantic.ckan.net/group/?group=http://ckan.net/group/llid>, as of 2011/09/06

3. What is Ex Libris supposed to do to promote Linked Open Library Data?

Provide open interfaces for legacy and future systems

- **Provide streaming interface** (event driven => real time) for all library data
- **Use open standards** (like XMPP) for interfaces

Enable Open Data

- Don't use ,open' as marketing term but in the clear sense of open definition
- Alma in general:
 - Provide possibility to **export all data** (e.g. statistics, user data) **in open formats**
- Alma's community zone:
 - **Openly licensed & accessible** as a whole **by everyone**

SWIB11

- **"Semantic Web in Libraries"** (SWIB)
hosted by hbz and ZBW – Leibniz
Information Centre for Economics
- Hamburg, **28. - 30.11.2011**
- Web stream will be available
- <http://swib.org/>

Thanks.

Questions?

With pleasure now/here or also later to
semweb@hbz-nrw.de