Using DKAN\textsuperscript{a} - an Open Source Portal Solution for Publishing Social Science Data: Lessons learnt at DSZ-BO

Johanna Vompras, University Library Bielefeld, johanna.vompras@uni-bielefeld.de

\textsuperscript{a}https://www.drupal.org/project/dkan

Motivation

In the past few years, more and more institutions, universities, and single researchers aim to provide their data for reuse to the public. From the perspective of universities, there is a huge need for rolling-out data portals across all disciplines – to enable researchers to document, collect, normalize data, and first of all:

to give them simple tools for presenting their research output.

One promising open source solution is DKAN, a Drupal implementation of the popular CKAN data publishing platform. It bundles

- Drupal core,
- two content types for representing datasets and resources,
- efficient data retrieval with adjustable faceted search.

Features

- customizable metadata schemas
- metadata-APIs
- and dataset harvesting capabilities

Flexibility: all Drupal modules are available, like bibliography, backup management, data visualization etc.

Screenshots

Key Facts

- Migration of the former (based on Javascript Framework EXTJS), BaseX database
- Cleaning, simplification and data transformation routines to import XML data (studies as DDI Instances)
- XML import into DKAN by the RSS Feeds Module
- Mapping definition DDI elements to fields of content type "Study"

References

- RDM at University Bielefeld: http://data.uni-bielefeld.de
- DRUPAL - Open Source CMS: https://www.drupal.org/
- Data Service Center for Business and Organisational Data: dszbo.uni-bielefeld.de/studienportal