

ZBW Labs: Publish Projects as Linked Data

Joachim Neubert
ZBW – Leibniz Information
Centre for Economics,
Germany
j.neubert@zbw.eu

Keywords: project repository; semantic publishing; linked data; Drupal; HTML5; RDFa1.1; doap; dcterms; schema.org

ZBW Labs¹ gives insight into software developments of the German National Library of Economics (ZBW). Maintained by the IIP (Innovative Information Systems and Publishing Technologies) department of the ZBW, it presents early prototypes as well as case studies in data visualization or full-fledged beta applications, which are not yet in full production state.

The use of a content management system (CMS) was primarily driven by the need for better collaborative editing of the German/English content of the web site, for better ways of user interaction, such as on-page-comments, and for improved internal and external linking.

Yet, we also wanted to bring the labs projects into the Web of Data—especially as some of these projects are Linked Data projects themselves. That implied providing an URI for each project, and offering additional (machine-readable, RDF) information under this URI. The

The screenshot shows the ZBW Labs website. At the top left is the ZBW logo (Leibniz-Informationszentrum Wirtschaft / Leibniz Information Centre for Economics) and the LABS logo (a blue flask with bubbles). Navigation links for 'News', 'Projects', and 'About us' are visible. A 'Deutsch' link is on the right. A word cloud on the left includes terms like 'Publishing Technologies', 'Visualizing Cultural Heritage', 'Linked Open Data', 'Web 2.0', 'Repositories', 'Automatic Indexing', 'Open Access', and 'Document analysis'. A search bar and login fields (Username, Password) are present. The main content area features a project titled 'Web Services for Economics'. To the left of the text is a small table with RDF triples:

asset
Asset-backed security
Asset
Asset market → Financial
Asset market approach
Asset pricing → Capital

The text describes the project as experimental thesaurus web services for humans and machines, designed to support query expansion in information retrieval applications. It notes that data originates from third-party datasets and follows REST design principles. Below the text is a 'Project Details' section with the following information:

- Name: Web Services for Economics
- Short Description: RESTful Semantic Web based terminology services
- Homepage: <http://zbw.eu/beta/econ-ws>
- Developer: Joachim Neubert
- Categories: API, Linked Open Data
- Created: 2009-03
- Project Status: Beta

At the bottom, there are logos for Public Domain, Drupal, W3C, and RDFa, along with the text '(rdf/xml, turtle, nt) Imprint'.

FIG. 1. HTML5/RDFa example page

¹ <http://zbw.eu/labs>

information ranges from project name and description, start and end dates, developers and funders (if any), to project status and category. Linking to other projects as well as to blog entries about certain aspects or experiences was essential. (Fig. 1) Linking out the other Linked Data entities (e.g., DBpedia) is still on the agenda.

To achieve these goals, the new labs site was built on Drupal 7. As a modern content management system, Drupal supports blogging, comments and other forms of user interaction, versioning and an editorial workflow. But most importantly, it allows definition of custom content types, to which newly defined fields can be attached (so project descriptions can have different fields from blog entries). Moreover, it is a Drupal 7 core functionality to support the creation of RDF. Classes can be assigned to the content types, and data and object properties to the fields. Drupal embeds this data into the generated HTML (in our case HTML5, produced by the Zen theme, and RDFa 1.1 compliant markup). As RDF vocabularies, dcterms, doap (description of a project), and schema.org were used. RDFa itself as well as the Drupal RDF mapping interface allow lists of attributes, which can be mixed from different vocabularies. This facilitated the semi-redundant combination of precise terms from the appropriate special vocabulary together with the more general dcterms properties and search-engine-friendly schema.org attributes.

These definitions can be accomplished by site builders in a very simple way through the administrative user interface (Fig. 2). Without the need of programming skills, or manipulating template files, the appropriate markup is produced and inserted by the Drupal “theming” subsystem. For the default entities of the Drupal system – articles, vocabularies and terms, users, comments – reasonable default mappings to skos, dcterms, foaf and scioc classes and properties are already defined.

DEVELOPER RDF MAPPING

RDF Predicates

schema:author, doap:developer, dc:creator

Enter a comma-separated list of predicates for *Developer* using CURIE syntax. For example: *foaf:familyName, foaf:lastName*

Attribute Type

rel

For fields containing literals—things such as plain text, html, numbers—use the **property** attribute. For fields containing references to other things—urls and node references, for example—use the **rel** or **rev** attribute.

FIG. 2. Assignment of multiple RDF mappings in the Drupal 7 admin user interface

To achieve Linked-Data-compliant “cool” URIs however took additional effort – especially because multi-lingual content had to be addressed. Another shortcoming showed up for nested field structures within one page, which are currently not supported by Drupal RDF and required custom theming functions. Also, RDFa markup for multiple custom entities within a given context (which for example would be necessary for the modeling of OAI-ORE aggregations) isn’t provided yet.

The overall experience however showed that semantically enhanced publishing has become available through a mainstream CMS tool such as Drupal, which hopefully will be enhanced with the upcoming versions.