

Linked Open Data University of Münster – Infrastructure and Applications

Carsten Kessler and Tomi Kauppinen

Institute for Geoinformatics, University of Münster, Germany
`carsten.kessler|tomi.kauppinen@uni-muenster.de`

Abstract. The Linked Open Data University of Münster (LODUM) project establishes a university-wide infrastructure to publish university data as Linked Open Data. The main goals are to increase visibility and accessibility of data produced and collected at the university, and to facilitate effective reuse of these data. This includes the goal to ease the development of applications and mashups based on the data, so that the common user can benefit from the LODUM data. This demo shows the LODUM infrastructure that facilitates application development, and two applications that demonstrate the potential of the LODUM data API.

1 The LODUM Project

Today’s research results are no longer limited to papers and books, but also include a variety of data, models, and software. Preserving and accessibility of these contents is fundamental to ensure the transparency and reproducibility of studies. The different scientific resources, be they publications, datasets, methods or tools should be annotated, interlinked and openly shared in order to make science more transparent and reproducible according to the Linked Science¹ approach [1]. Likewise, developers and endusers can benefit from administrative data published as Linked Open Data.

The LODUM project tackles this long-term goal by implementing a strategy that aims to improve the transparency and visibility of the university, publishing any non-sensitive data online following the Linked Data principles. Data sources existing across the different sites of the university remain in place, leaving the control and responsibility in the hands of the original owner. Such data can be linked to and accessed from <http://data.uni-muenster.de>. This covers both scientific data and publications, as well as administrative data such as building databases and course schedules. LODUM is laid out as a long-term strategy that will open up and connect different data sources across the 15 faculties and departments step-by-step. With LODUM, the University of Münster the first German university to implement such a program, following the early examples from the UK.² The growing number of publications on Linked Data in science and education shows that this approach is gaining momentum (see, e.g., [2,3,4]).

¹ See <http://linkedsience.org>.

² See <http://linkeduniversities.org/lu/index.php/datasets-and-endpoints/>.

This demo shows the technical infrastructure to manage data and make them available online. Moreover, we introduce two sample applications that have been built on the LODUM infrastructure: a 3D productivity map of the university, and the WWU App, a Web application optimized for mobile phones that provides information about the university for students, staff, and visitors.

2 Infrastructure

This section documents the LODUM infrastructure, i.e., the technical workflow and the vocabularies in use.

2.1 Technical Workflow

The LODUM infrastructure consists of three layers, as shown in Figure 1. At the core of this setup, an OWLIM triple store hosts the LODUM data. The management tools support tasks such as vocabulary management and link discovery between the initially disconnected datasets imported into LODUM. The retrieval tools expose the LODUM via type-specific HTML pages as well as raw RDF data, which can also be queried via the SPARQL endpoint at <http://data.uni-muenster.de/sparql>. In order to get data from existing systems into this infrastructure, we have developed custom triple factories that fetch the data from the different existing systems, convert them to RDF, and push them to the triple store.

2.2 Vocabularies in Use

We reuse existing vocabularies as much as possible to provide semantic annotations of the data we offer in LODUM. In addition we also created new needed vocabularies to introduce new classes and properties.

The Bibliographic Ontology (BIBO) is used for all bibliographic resources such as books and articles. The Dublin Core Metadata Element Set is used for all occasions where basic metadata are required. The Friend of a Friend (FOAF) vocabulary is used for all information about people and organizations. The W3C Basic Geo Vocabulary is used for all things that are georeference via WGS84 lat/long coordinates. The Teaching Core Vocabulary (TEACH) was used to annotate course descriptions. Publishing Requirements for Industry Standard Metadata (PRISM) provided additional properties for annotating publishing content.³

³ See the following URLs for these vocabularies: <http://bibliontology.com/>, <http://dublincore.org/documents/dces/>, <http://www.foaf-project.org/>, <http://www.w3.org/2003/01/geo/>, <http://linkedscience.org/teach/ns/>, and <http://www.idealliance.org/specifications/prism/>.

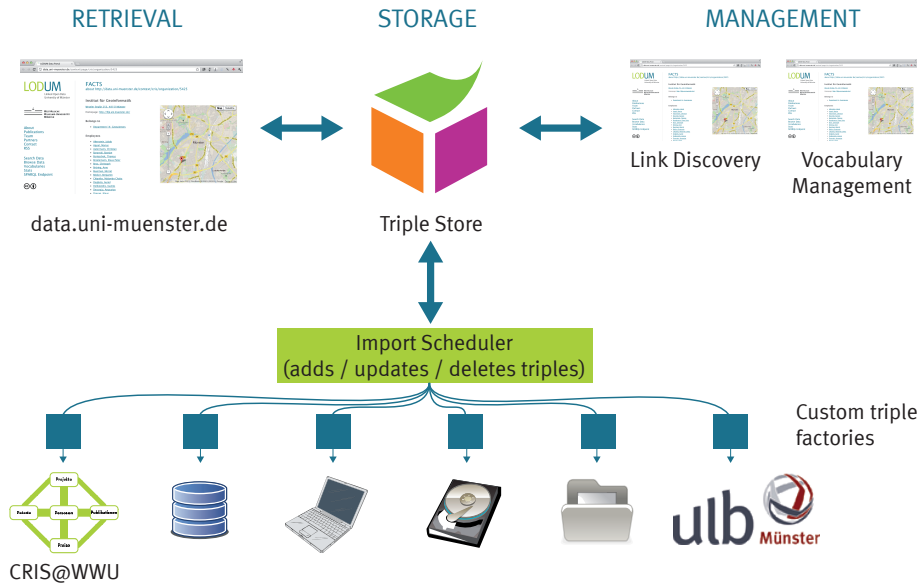


Fig. 1. High-level overview of the LODUM architecture.

3 Applications

This section introduces the two applications we would like to demonstrate, the LODUM productivity map and the WWU app.

3.1 Productivity Map of University Researchers

The productivity map for Google Earth⁴ shown in Figure 2 is an example of how LODUM facilitates data analysis. It renders the university buildings in 3D, where the building height indicates the number of publications written by researchers working in the respective building. The geometries of the buildings are stored according to the emergent GeoSPARQL standard [5]. This interactive map can act as an entry point for the evaluation of and comparison between different departments.

For the building height, the absolute number of papers is normalized by the number of researchers working in the given building for a more balanced impression. The buildings are split in two parts: the lower part indicates the number of journal papers, whereas the upper part represents all other publications. Clicking either of these two parts opens a pop-up with the actual numbers. The distribution of publications between the different institutions in a building is visualized as a pie chart. The pop-ups also include links to the SPARQL queries to pull the

⁴ See <http://data.uni-muenster.de/context/static/lodum-productivity-map.kml> for the KML file and <http://youtu.be/QvQ2GF2Zz5g> for a video tour.

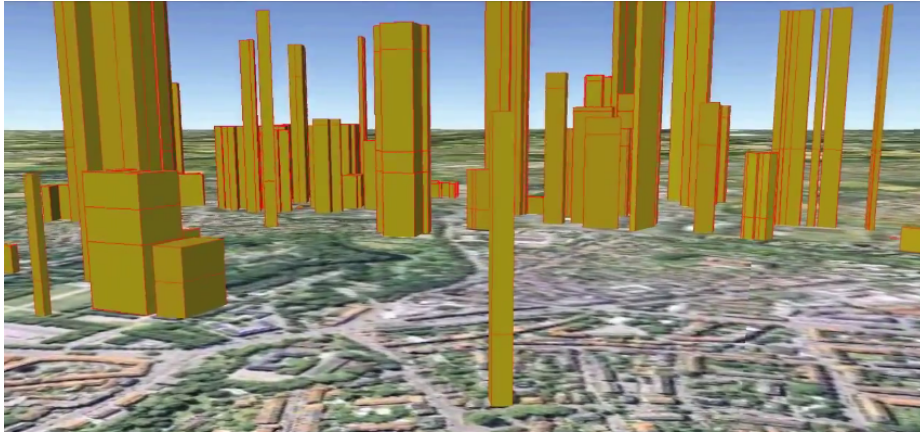


Fig. 2. Screen shot of the map in Google Earth.

data for the given building out of our store, so that interested developers can learn how we built this map.

3.2 WWU App

The *Westfälische Wilhelms-Universität Münster* (WWU) wants to offer students, staff and visitors information about cafeteria menus, navigation instructions to university buildings, and an overview map of all university buildings on their mobile phone. The latest version of this *WWU App*, accessible at <http://app.lodum.de> is a platform-independent Web app based on the LODUM data store, as shown in Figure 3. It has been developed to replace existing native apps that had the data *baked in*, thus requiring frequent updates, and that were difficult to maintain for the wide range of different platforms. The initial version of the new Web app has been developed in a seminar taught by the authors. It has been built around the PrimeFaces JSF framework and a Glassfish application server that pulls the data from the SPARQL endpoint.

4 Conclusions

The University of Münster, Germany, is one of the first universities to commit to a institution-wide Linked Open Data program. The goal of the LODUM initiative is to increase transparency and comprehensibility both for research and for administrative matters. Beyond the technical infrastructure required to put this undertaking into practice—server facilities to run a triple store, SPARQL endpoint, data dumps, backups, and synchronization jobs—fostering a change in the mindset of the university’s research community is key for the success of LODUM. In this paper we described two applications demonstrating the use of the LODUM infrastructure. With these demonstrations we want to show that

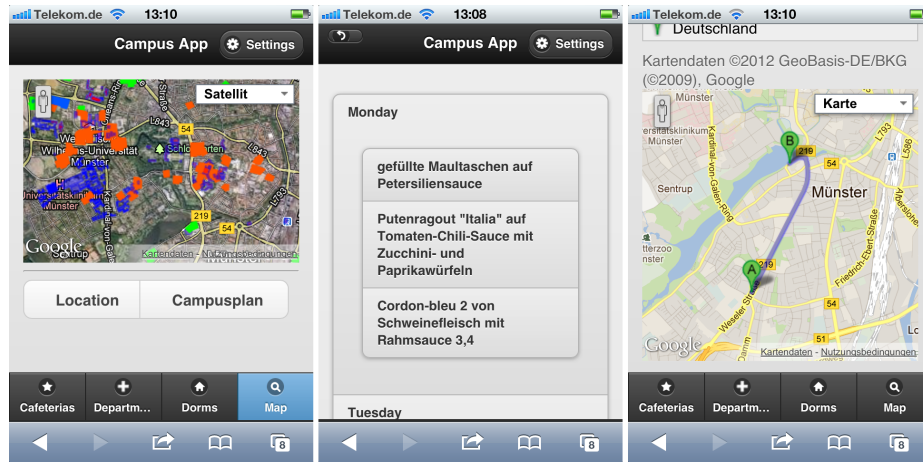


Fig. 3. Screen shots with overview of university buildings (left), cafeteria menus (middle), and routing to a cafeteria (right).

the real potential of Linked Data is in the ease of developing applications on top of the data using a standardized API—i.e. the way to access data online.

Acknowledgements

This research has been funded by the *Linked Open Data University of Münster* project (<http://lodum.de>) and the International Research Training Group *Semantic Integration of Geospatial Information* (DFG GRK 1498). We would like to thank Steffan Voss and Christopher Stephan for developing the WWU App.

References

1. Kauppinen, T., Baglatzi, A., Keßler, C.: Linked Science: Interconnecting Scientific Assets. In Critchlow, T., Kleese-Van Dam, K., eds.: Data Intensive Science. CRC Press, USA (forthcoming 2012)
2. Shotton, D., Portwin, K., Klyne, G., Miles, A.: Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article. *PLoS Comput Biol* **5**(4) (2009)
3. Demartini, G., Enchev, I., Gapany, J., Cudré-Mauroux, P.: The Bowlogna Ontology: Fostering Open Curricula and Agile Knowledge Bases for Europe's Higher Education Landscape. *Semantic Web Journal* (accepted)
4. Zablith, F., d'Aquin, M., Brown, S., Green-Hughes, L.: Consuming linked data within a large educational organization. In: Second International Workshop on Consuming Linked Data (COLD) at 10th International Semantic Web Conference (ISWC 2011). (2011)
5. Open Geospatial Consortium: OGC GeoSPARQL – A Geographic Query Language for RDF Data. Request for comments available from <http://www.opengeospatial.org/standards/requests/80>; last accessed 2011-12-29 (2011)