

The *Catalogus Professorum Lipsiensis* – Semantics-based Collaboration and Exploration for Historians

Thomas Riechert*, Ulf Morgenstern+, Sören Auer*, Sebastian Tramp*, and Michael Martin*

* AKSW, Institut für Informatik, Universität Leipzig, Pf 100920, 04009 Leipzig
{lastname}@informatik.uni-leipzig.de, <http://aksw.org>
+ Historisches Seminar, Universität Leipzig, Pf 100920, 04009 Leipzig
{lastname}@uni-leipzig.de, <http://www.uni-leipzig.de/histsem>

1 Introduction

The World Wide Web (WWW), as an ubiquitous medium for publication and exchange, already significantly influenced the way how historians work: the availability of public catalogs and bibliographies enable efficient research of relevant content for a certain investigation; the increasing digitization of works from historical archives and libraries, in addition, enables historians to directly access historical sources remotely. The capabilities of the WWW as a medium for collaboration, however, are only starting to be explored. Many historical questions are only answerable by combining information from different sources, from different researchers and organizations. Furthermore, after analyzing original sources, the derived information is often more comprehensive than can be captured by simple keyword indexing.

In [3] we report about the application of an adaptive, semantics-based knowledge engineering approach for the development of a prosopographical knowledge base. In this demonstration we will showcase the comprehensive prosopographical knowledge base and its potential for applications. In prosopographical research, historians analyze common characteristics of historical groups by studying statistically relevant quantities of individual biographies. Untraceable periods of biographies can be determined on the basis of such accomplished analyses in combination with statistically examinations as well as patterns of relationships between individuals and their activities. In our case, researchers from the Historical Seminar at the University of Leipzig aimed at creating a prosopographical knowledge base about the life and work of professors in the 600 years history of University of Leipzig ranging from the year 1409 till 2009 - the *Catalogus Professorum Lipsiensis* (CPL).

2 Architectural Overview

The system architecture of CPL comprises a combination of different applications, which interact using standardized interfaces as illustrated in Figure 1. We divided the architecture into two separated zones (public and protected zone) due to technical constraints and in order to prevent security problems.

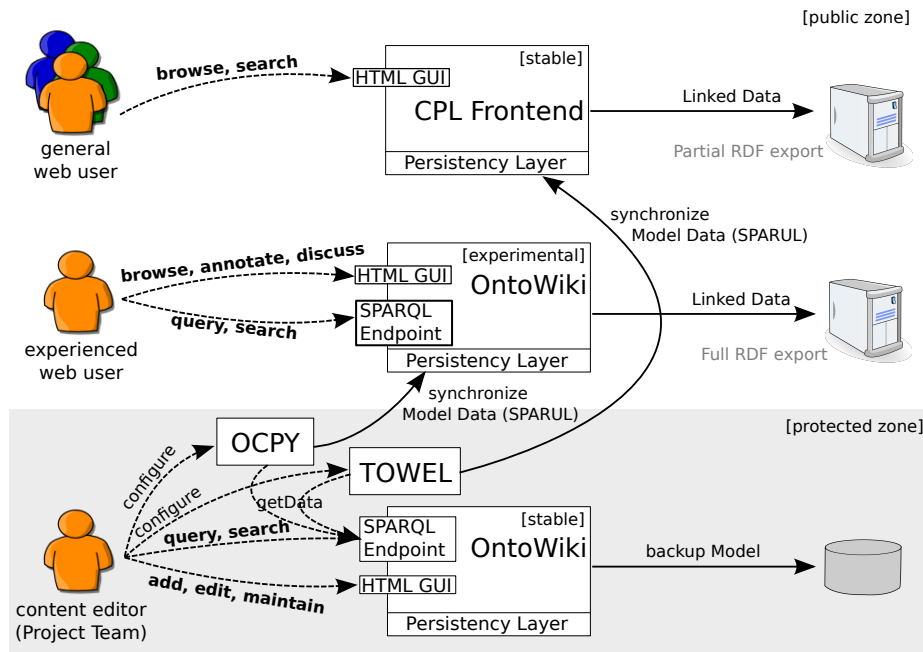


Fig. 1: Architectural overview about the project platforms

The semantic data wiki OntoWiki [1] located in the protected layer uses the *Catalogus Professorum Model*¹ (CPM), which comprises several ontologies and vocabularies for structuring the prosopographical information. The project team is working collaboratively and spatially distributed (e.g. in archives or libraries) to collect, structure and validate information about persons and institutions relevant to this knowledge domain.

For general web users the catalog is integrated in the public website of the University of Leipzig². A simplified user interface consisting of plain HTML and Linked Data³ resources is generated nightly from the knowledge base. The historians are able to interact with CPL via an experimental version⁴ of OntoWiki. The version of the catalog available there is synchronized with the protected OntoWiki installation, transforms the exported data considering any linked knowledge bases and imports the changed data into this experimental installation. The project platform is usable for humans accessing web interfaces for reading or editing data. OntoWiki also provides a number of generic access interfaces, which include SPARQL and Linked Data endpoint and a Semantic Pingback server. On top of these generic access interfaces application specific access interfaces are deployed. These are described in more detail in the remainder and will be presented in the demonstration.

¹ Available at: <http://catalogus-professorum.org/cpm/>

² <http://www.uni-leipzig.de/unigeschichte/professorenkatalog/>

³ <http://www.w3.org/DesignIssues/LinkedData.html>

⁴ <http://catalogus-professorum.org/>

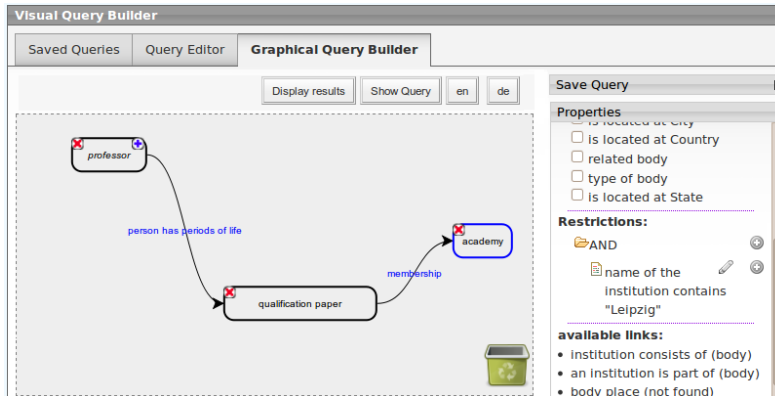


Fig. 2: *Visual Query Builder*.

3 Public CPL website and Linked Data

CPL is not just a tool for historians, but aims to showcase the results of historic research to the wider general public. For that purpose a special public website was created. The user interface of the public website is geared towards simplicity. The knowledge base can be explored by epochs, faculties, functions of professors (i.e. rector or dean) or alphabetically. Professors of the day are automatically selected based on important days in the life of a professor (i.e. birth or death). Furthermore, the public website comprises a full-text search, which searches within all literals stored in the CPL knowledge base.

Using OntoWiki's build-in endpoint functionality CPL is immediately available as Linked Data. Within the Linking Open Data effort, hundreds of data sets have already been connected to each other via `owl:sameAs` links. By interlinking CPL with other related datasets, we aim at establishing CPL as a linked data crystallization point for academic prosopographical knowledge.

4 Visual Query Builder.

OntoWiki also serves as a SPARQL endpoint, however, it quickly turned out that formulating SPARQL queries is too tedious for the historian domain experts. In order to simplify the creation of queries for the historians, we developed the *Visual Query Builder*⁵ (VQB) as an OntoWiki extension, which is implemented in JavaScript and communicates with the triple store using the SPARQL language and protocol. VQB allows to visually create queries to the stored knowledge base and supports domain experts with an intuitive visual representation of query and data. Developed queries can be stored and added via drag-and-drop to the current query. This enables the reuse of existing queries as building blocks for more complex ones. VQB also supports the *set-based browsing* paradigm by visualizing different connectives, such as join, union,

⁵ <http://aksw.org/Projects/OntoWiki/Extension/VQB>

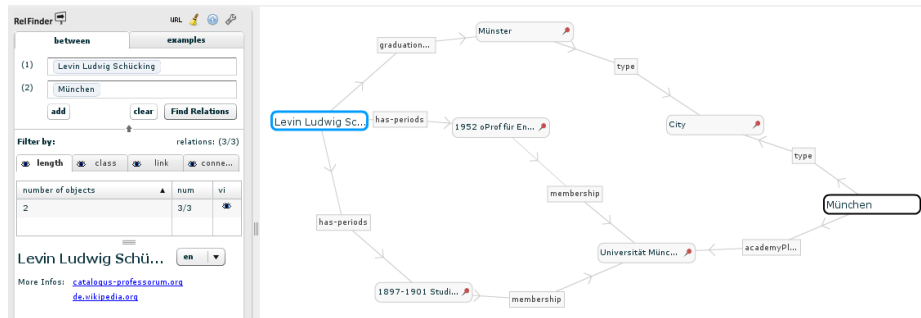


Fig. 3: Visualization of relationships in *RelFinder*.

intersection, difference between queries. The incremental query building is facilitated by displaying results already during query creation. The VQB user interface is visualized in Figure 2. The user interface can be adjusted by scaling or deactivating unused panels.

5 Relationship Finder.

An important aspect of historical investigations is the search for relationships between different persons or entities of interest. An application supporting such investigations within RDF knowledge knowledge bases is *RelFinder*⁶ [2]. With the help of *RelFinder* relationships between individual entities can be easily discovered and visualized. Figure 3, for example, visualizes the relationship between the entities *Schücking* and *München*⁷. In this example, three connections were found and visualized as paths through the knowledge base. *RelFinder* is a generic tool and can be used in conjunction with arbitrary SPARQL endpoints.

References

1. Sören Auer, Sebastian Dietzold, and Thomas Riechert. OntoWiki – A Tool for Social, Semantic Collaboration. In *Proceedings of the 5th International Semantic Web Conference, ISWC2006*, Athens, GA, USA, 2006.
2. Philipp Heim, Sebastian Hellmann, Jens Lehmann, Steffen Lohmann, and Timo Stegemann. RelFinder: revealing relationships in RDF knowledge bases. In *Proceedings of the 4th International Conference on Semantic and Digital Media Technologies*, 5887 of LNCS, pages 182–187. Springer, 2009.
3. Thomas Riechert, Ulf Morgenstern, Sören Auer, Sebastian Tramp, and Michael Martin. Knowledge engineering for historians on the example of the *catalogus professorum lipsiensis*. In *Proceedings of the 9th International Semantic Web Conference, ISWC2010*, Shanghai, China, 2010.

⁶ Online at: <http://relfinder.semanticweb.org>

⁷ More interesting relationships obtained from RelFinder are listed at: <http://catalogus-professorum.org/tools/relfinder>