

Linked Data Query Wizard: A Tabular Interface for the Semantic Web

Patrick Hoeffler¹, Michael Granitzer²,
Vedran Sabol¹, and Stefanie Lindstaedt¹

¹ Know-Center, Graz, Austria

² University of Passau, Germany

{phoeffler,vsabol,slind}@know-center.at
michael.granitzer@uni-passau.de

Abstract. Linked Data has become an essential part of the Semantic Web. A lot of Linked Data is already available in the Linked Open Data cloud, which keeps growing due to an influx of new data from research and open government activities. However, it is still quite difficult to access this wealth of semantically enriched data directly without having in-depth knowledge about SPARQL and related semantic technologies. In this paper, we present the Linked Data Query Wizard, a prototype that provides a Linked Data interface for non-expert users, focusing on keyword search as an entry point and a tabular interface providing simple functionality for filtering and exploration.

Keywords: linked data, interfaces, semantic web, sparql, rdf

1 Introduction

The Linked Open Data cloud provides an impressive wealth of semantically enriched, openly available Linked Data. However, this Linked Data is basically only accessible for experts in semantic technologies who know how to write SPARQL queries. And even for those who know how to use SPARQL, it can be quite laborious at times, especially while trying to explore an unknown SPARQL endpoint.

Therefore, the goal of the presented prototype – the Linked Data Query Wizard¹ – is to provide an easy-to-use interface for accessing Linked Data. It should be suitable for non-expert users without any prior knowledge of SPARQL or other semantic technologies.

The working hypothesis for the Linked Data Query Wizard is: There are not that many people who speak SPARQL and are familiar with graph structures. On the other hand, many people know spreadsheet applications like Microsoft Excel. Therefore, the idea is to develop a web-based tool that brings the graph structure of Linked Data into tabular form (see figure 1) and provides easy-to-use interaction possibilities for filtering and exploring Linked Data by using metaphors and techniques the users already know.

¹ <http://code.know-center.tugraz.at/search>

CODE Linked Data Query Wizard

Label [graz] x	Description	Type Funding	Partner 999977948	PartnerRole	Amount	Add column...
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project MATURE		Funding	999977948	Partner	682950	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project CopPer		Funding	999977948	Partner	525452	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project SECO		Funding	999977948	Partner	564378	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project ECRYPT II		Funding	999977948	Partner	132000	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project MOGENES		Funding	999977948	Partner	445000	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project COCONUT		Funding	999977948	Partner	296716	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project IMPACT		Funding	999977948	Partner	691982	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project ICT-ENSURE		Funding	999977948	Coordinator	668886	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project PRE-DRIVE		Funding	999977948	Partner	62836	
Funding of partner TECHNISCHE UNIVERSITAET GRAZ in project CHOSEN		Funding	999977948	Partner	256620	

Displaying 10 of 41 results

[Load more results ...](#)

This is CODEresearch in progress.

Fig. 1. Current results page of the Linked Data Query Wizard.

2 Research Context

The Linked Data Query Wizard is part of the CODE project [1], a research project funded by the European Union. As described in [2], the vision of CODE is to establish a sophisticated ecosystem for Linked Data. The extraction of knowledge encapsulated in scientific research papers along with its release as Linked Data serves as the major use case. A web-based visual analytics interface should empower end users to analyze, integrate, and organize data.

Our prototype is part of CODE’s Visual Analytics work package [3]. The goal of the work package is to develop a web-based visual analytics platform that enables non-expert users to engage in a visually supported, collaborative analysis of Linked Data, and the Linked Data Query Wizard will play a crucial role in this undertaking.

3 Related Work

Although the Semantic Web has matured in recent years, and semantic technologies have become quite powerful, the Linked Open Data cloud is still only accessible for semantic technology experts and programmers. The problem of easy-to-use interfaces for accessing Linked Data is still largely unsolved. The majority of current tools are not aimed at non-expert users. As an example, the popular Semantic Web search engine Sindice [4] is practically unusable for people without a deep understanding of semantic technologies.

Currently, only very few web-based tools use tables for representing Linked Data. One such example would be Freebase Parallax [5]. Although its main feature is the ability to browse sets of related things, it also provides a table view for these result sets. Another online tool that shares similarities with our prototype is the Falcons Explorer [6]. Both tools feature a search box as the main entry point – an idea that is also central to our prototype. However, in

both tools, the table view is not the central focus, and usability is hindered by a multitude of different viewing options that are bound to confuse inexperienced users.

Another tool that shares a few similarities with our prototype is OpenRefine [7] (formerly known as Google Refine and Freebase Gridworks). It supports RDF, and there are also extensions such as LODRefine [8] that focus on Linked Data – however, OpenRefine’s main focus is cleaning up tabular data, and it’s also not available as a web service, even though its main interface is browser-based.

Our prototype also supports the current working draft of W3C’s RDF Data Cube Vocabulary [9] which provides a semantic framework for expressing statistical datasets as Linked Data. Datasets that comply with the RDF Data Cube standard can easily be displayed, filtered, and explored using the Linked Data Query Wizard. Because this standard is still very young, to the best of our knowledge, there are currently no other tools – with a feature set similar to our prototype – that support RDF Data Cubes.

4 Key Technologies

The Linked Data Query Wizard is a purely web-based system. It uses Python 2.7 on the back end and HTML5, CSS3, and JavaScript – compiled from CoffeeScript – on the front end. Its main SPARQL endpoint is provided by a bigdata [10] RDF database, which also provides an integrated full-text search [11].

Sadly, full-text search is sorely lacking from the current SPARQL specification [12], which is why certain SPARQL endpoints have come up with workaround solutions. Therefore, only Virtuoso and bigdata are currently supported as SPARQL endpoints by the Linked Data Query Wizard. However, since one of our prototype’s design philosophies is to use Semantic Web standards such as SPARQL wherever possible, support for other suitable endpoints could be added with minimal effort at a later point.

The Linked Data Query Wizard also makes use of certain SPARQL 1.1 features, especially the aggregation functions. The `COUNT()` function is critical and already in use by the current prototype for displaying the number of results for a given query. It is also planned to extend the prototype in order to let users calculate simple aggregate functions for a given dataset, using functions such as `SUM()` and `AVG()`. For advanced filtering mechanisms, functions such as `MIN()` and `MAX()` might be needed as well.

5 Initial Results

The Linked Data Query Wizard is currently available online as an early beta version. In its current form, it offers two entry points: Users can either initiate a keyword search, or they can select any available dataset, represented as an RDF Data Cube (see figure 2). In both cases, the users get presented with the results in tabular form, similar to what they are used from spreadsheet applications.

They can choose which columns (i.e. RDF predicates) they are interested in, and they can set filters to narrow down the displayed data.

Though the current functionality of the prototype is still rather limited, first usage experiments have shown that the tool can be helpful in exploring the data and respective data structures of unknown SPARQL endpoints. For example, with one keyword search and less than 20 clicks, it is possible to either create a dataset of all coordinators of finished EU FP7 research projects and their received funding, or a list of public companies located in the US and their number of employees.

Additionally, the Linked Data Query Wizard has been under permanent scrutiny of fellow researchers from the CODE project for several weeks now, providing us with valuable feedback on stability and usability issues as well as helpful feature requests.

CODE Linked Data Query Wizard

Search bar:

Buttons: Search EU Open Data | Search CODE

Currently available data sets (provided by EU Open Data):

- % of basic public services for citizens, which are fully available online
- % of basic public services for enterprises, which are fully available online
- % of disadvantaged people who are regular internet users (at least once a week)
- % of distinct organisations participating in FP7-ICT projects for the first time
- % of enterprises exchanging business documents suitable for automatic processing
- % of enterprises interacting online with public authorities
- % of enterprises providing portable devices to some of their persons employed
- % of enterprises returning filled in forms to public authorities, over the internet
- % of enterprises sending or receiving e-invoices in a format suitable for automatic processing

This is CODEResearch in progress.

Fig. 2. Current entry page of the Linked Data Query Wizard.

6 Conclusions and Future Work

In this paper, we presented the Linked Data Query Wizard, a tabular approach for filtering and exploring Linked Data.

The next steps regarding the prototype will be to expand its functionality, focusing on better filter mechanisms as well as more advanced exploration features that incorporate the underlying semantic structure. Additionally, the Linked Data Query Wizard will be integrated with a tool for visualizing Linked Data as well as a semantic enrichment service for turning generic RDF into RDF Data Cubes.

The development of the prototype will continue throughout the rest of the year, leading to a final evaluation at the beginning of 2014.

Acknowledgments The Linked Data Query Wizard is being developed within the CODE project at the Know-Center, Graz, Austria. The CODE project is funded by the EU Seventh Framework Programme, grant agreement number 296150. The Know-Center is funded within the Austrian COMET Program – Competence Centers for Excellent Technologies – under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth, and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency (FFG).

References

1. CODE: Commercially Empowered Linked Open Data Ecosystems in Research, <http://code-research.eu/>
2. Stegmaier, F., Seifert, C., Kern, R., Hoefler, P., Bayerl, S., Granitzer, M., Kosch, H., et al.: Unleashing Semantics of Research Data. In: Proceedings of the Second Workshop on Big Data Benchmarking (WBDB2012.in). Pune, India (2012)
3. CODE: Visual Analytics, <http://code-research.eu/visual-analytics>
4. Tummarello, G., Delbru, R., Oren, E.: Sindice.com: Weaving the open linked data. The Semantic Web, 552-565. Springer Berlin Heidelberg (2007)
5. Huynh, D. F., Karger, D. R.: Parallax and Companion: Set-based Browsing for the Data Web. WWW Conference, ACM, 2009.
6. Cheng, G., Wu, H., Gong, S., Ge, W., Qu, Y.: Falcons Explorer: Tabular and Relational End-user Programming for the Web of Data. Semantic Web Challenge, 2010.
7. OpenRefine, <https://github.com/OpenRefine>
8. LODRefine, <http://code.zemanta.com/sparkica/>
9. W3C: RDF Data Cube Vocabulary, <http://www.w3.org/TR/vocab-data-cube/>
10. bigdata, <http://www.systap.com/bigdata.htm>
11. bigdata Full Text Search, <http://sourceforge.net/apps/mediawiki/bigdata/?title=FullTextSearch>
12. W3C: SPARQL 1.1 Query Language, <http://www.w3.org/TR/sparql11-query/>