Query Wizard and Visualization Wizard

Easy to use Search, Exploration and Visual Analysis of Linked Data

Although the concept of Linked Data has been increasing in popularity, easy-to-use interfaces to access and make sense of the actual data are still few and far between. The CODE project's Query Wizard and Vis Wizard are here to help.

The amount of Linked Data available on the Web keeps growing, mainly 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 of semantic technologies.

Therefore, one of the goals of the **EU-funded CODE project** has been to develop a web-based visual analytics platform that enables non-expert users to easily perform exploration and analysis tasks on Linked Data. The vision of CODE is to establish a tool chain for the extraction of knowledge encapsulated in scientific research papers along with its release as Linked Data. A web-based visual analytics interface should empower end users to analyze, integrate, and organize the data. The CODE Query Wizard and the CODE Vis Wizard play a crucial role in this undertaking.

When it comes to working with data, many people know how to use spreadsheet applications like Microsoft Excel. In comparison, there are very few people who know SPARQL, the W3C standard language to query Linked Data. The **CODE Query Wizard** provides a web-based interface that dramatically **simplifies the process of displaying, accessing, filtering, exploring, and navigating the Linked Data** that's available through a SPARQL endpoint. The main

Figure 1: An RDF Data Cube provided by the European Open Data Portal is displayed and filtered in the CODE Query Wizard.

innovation of the interface is that it **turns the graph structure of Linked Data into tabular form and provides easy-to-use interaction possibilities** by using metaphors and techniques that the end users are already familiar with.
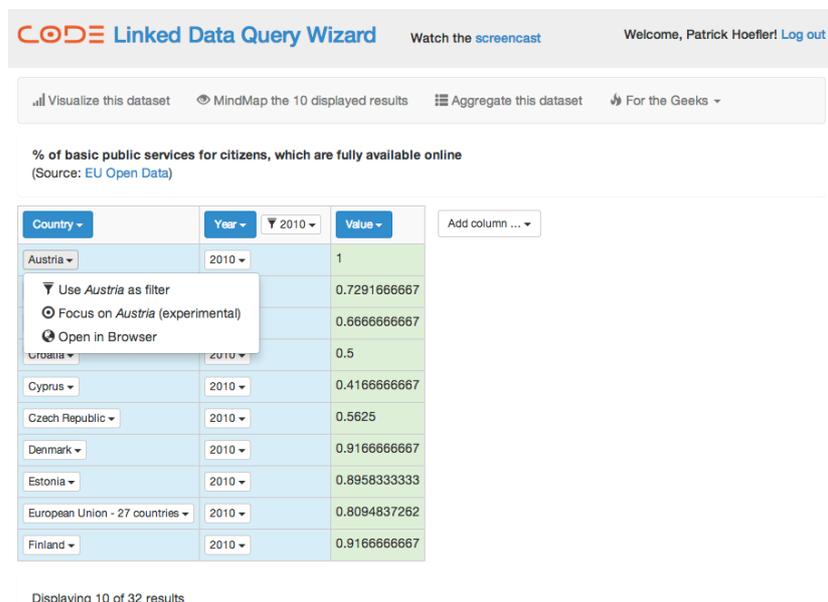
The CODE Query Wizard offers two entry points: Users can either initiate a **keyword search over a Linked Data repository**, or they can select any of the **already available datasets, represented as RDF Data Cubes**. In both cases, the CODE Query Wizard presents a table containing the results. The users can then choose which columns they are interested in, and they can set filters to narrow down the

displayed data. Additionally, they can explore the data by "focusing" on an entity, or they can aggregate a dataset to get a summary of the data.

Once the users are happy with their selected data, they can visualize it using the **CODE Vis Wizard**. This tool enables **visual analysis of Linked Data and supports the user by automating the visualization process**. This means that after analyzing the structural and semantic characteristics of the provided Linked Data, the CODE Vis Wizard automatically suggests any of the 10 currently available visualizations that are suitable for the provided data. Furthermore, the Vis Wizard **automatically maps the data on the available visual channels of the chosen visualization**. If the users wish to adjust the mapping, they can do so with a few simple clicks.
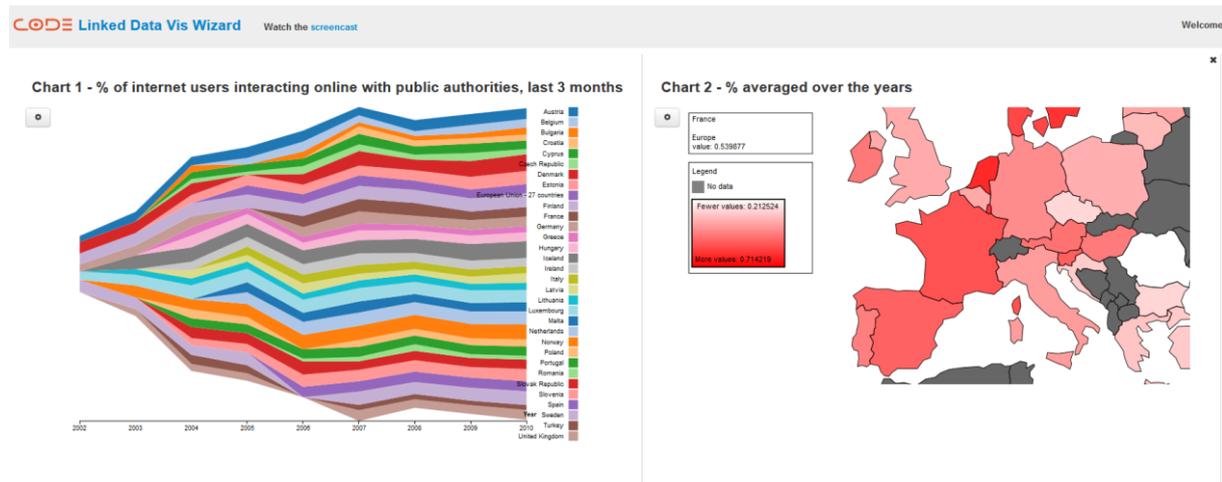

Figure 2: The CODE Vis Wizard displays an interactive visual representation of the percentage of users interacting online with authorities for each country. Shown on left is the trend over the years, on right - as a result of data aggregation - the average over years for each country.

Usually, more than one visualization is suitable for any given dataset. In that case, multiple **visualizations can be displayed side by side**. When certain parts of the data are selected in on of the visualizations, they are automatically highlighted in all of the others as well. This can provide **quick insights into complicated data**, taking advantage of the powerful human visual perception system.

The CODE Query Wizard and Vis Wizard are purely web-based systems. They currently support Virtuoso, OWLIM and bigdata SPARQL endpoints, since all of those also provide integrated full-text search. However, since the prototypes have been designed to use Semantic Web standards such as SPARQL wherever possible, support for other suitable endpoints could be added at a later point with minimal effort.

Both prototypes have been developed within the CODE project (funded by the EU Seventh Framework Programme, grant agreement no. 296150) at the Know-Center (http://www.know-center.at) in Graz, Austria, with support by their project partners University of Passau, Mendeley (London) and MeisterLabs (Vienna). The project has started in May 2012 and will finish in April 2014.

Useful Links:

- CODE Project Website: http://code-research.eu/
- CODE Query Wizard: http://code.know-center.tugraz.at/search
- CODE Visualization Wizard: http://code.know-center.tugraz.at/vis
- CODE Publications: http://code-research.eu/publications