Evaluation of LinDA and Comparison with other tools for visualization of Linked Data

List of Alternate Tools Specification

Lab Semantic Web Technologies – WS14/15

Authors:

Rose-Mary Owusuaa Mensah

Nishananth Baskaran

Ugochukwu Chimbo Ejikeme

Nikhil Patra

Mentor:

Dr. Fabrizio Orlandi

Table of Contents

1. Introduction

1.1. Revision History

Edition	Date	Authors	Remarks
01	11.24.2014	Rose-Mary Owusuaa Mensah, Nishananth Baskaran, Ugochukwu Chimbo Ejikeme, Nikhil Patra	Creation

1.1. Purpose

The purpose of this document is to provide a list of alternate tools for the visualization of RDF data sets. This will be helpful in the comparison of LinDA with the possible tools from the list. Each tool has its own purpose and provides its own features. This comparison will help LinDA to identify its own pros and cons and how competitive compared to the other tools.

This document is intended for the Supervisors of our Semantic Web Technologies Lab at the University of Bonn. It can also be used by the stakeholders of Enterprise Information Systems such as software developers, managers and businesses.

1.2. Scope

The scope of this document is to provide the stakeholders a list of alternate tools like LinDA which can be used for visualization of datasets. This document gives the information about the tools like description, file formats supported, and the like.

1.3. Definitions, Acronyms, Abbreviations

- LinDA: Linked Data Analytics

- RDF: Resource Description framework is a family of World Wide Web Consortium (W3C) specifications used as a general method for conceptual description or modelling of information that is implemented in web resources, using a variety of syntax notations and data serialization formats.

- Evaluators: A group of people who are involved in understanding the product quality which are: usability, efficiency and scalability and compare the new tool with the already existing tool.

- SPARQL: SPARQL Protocol and RDF Query Language

1.4. References

Reference Name	Reference Link
Tableau Public	http://www.tableausoftware.com/public/faq
Cytoscape	http://www.cytoscape.org/what_is_cytoscape.html
OntoWiki CubeViz	http://cubeviz.aksw.org/
Prefuse	http://prefuse.org/doc/faq/#
gFacet	http://www.visualdataweb.org/gfacet.php
RDF:SynopsVIZ	http://www.w3.org/2001/sw/wiki/Rdf:SynopsViz
Sgvizler	http://dev.data2000.no/sgvizler/
Visualbox	http://alangrafu.github.io/visualbox/

2. Tools Description

2.1. Tableau Public

Tableau Public is a free service that lets anyone publish interactive data to the web. Once on the web, anyone can interact with the data, download it, or create their own visualizations of it. No programming skills are required. Be sure to look at the gallery to see some of the things people have been doing with it. Tableau Public includes a free desktop product that you can download and use to publish interactive data visualizations to the web. The Tableau Public desktop saves work to the Tableau Public web servers – nothing is saved locally on your computer. All data saved to Tableau Public will be accessible by everyone on the internet, so be sure to work only with publically available (and appropriate) data.

File Format Supported: Tableau Public can connect to Microsoft Excel, Microsoft Access, and multiple text file formats. It has a limit of 1,000,000 rows of data that is allowed in any single file.

Download Link: http://www.tableausoftware.com/public/download

2.2. Cytoscape

Cytoscape is an open source software platform for visualizing molecular interaction networks and biological pathways and integrating these networks with annotations, gene expression profiles and other state data. Although Cytoscape was originally designed for biological research, now it is a general platform for complex network analysis and visualization. Cytoscape core distribution provides a basic set of features for data integration, analysis, and visualization.

File Format Supported: Cytoscape can read file structures that are delimited Text or Excel files.

Download Link: http://www.cytoscape.org/download.php

2.3. OntoWiki CubeViz

CubeViz was developed based on OntoWikiwith the scope of offering user-friendly exploration possibilities for statistical data represented in RDF with the RDF DataCube vocabulary. CubeViz represents the statistical dataset to be visualized as a facetted based browsing component. This component enables the users to select interesting parts of the dataset. After selection the user can proceed while clicking on the button Update Selection / Update Chart. One can instantly start using CubeViz while clicking on the button Start CubeViz above. As a result CubeViz processes a chart according to user's selection. The current version of CubeViz processes basic chart types - such as line, barandpie chart facilitating the exploration of up to two statistical dimensions in a data structure.

File Format Supported: Microsoft Excel, CSV Files.

Download Link: It is an online facetted browser. http://cubeviz.aksw.org/

2.4. Prefuse

Prefuse is an extensible software framework for helping software developers create interactive information visualization applications using the Java programming language. It can be used to build standalone applications, visual components embedded in larger applications, and web applets. Prefuse intends to greatly simplify the processes of representing and efficiently handing data, mapping data to visual representations (e.g., through spatial position, size, shape, color, etc.), and interacting with the data. Some of the features of Prefuse include:

- Table, Graph, and Tree data structures supporting arbitrary data attributes, data indexing, and selection queries, all with an efficient memory footprint.
- Components for layout, color, size, and shape encodings, distortion techniques, animation, and more.
- A library of interaction controls for common interactive, direct-manipulation operations.
- Animation support through a general activity scheduling mechanism.
- View transformations supporting panning and zooming, including both geometric and semantic zooming.
- Dynamic queries for interactive filtering of data, integrated text search.
- A built in, SQL-like expression language for writing queries to prefuse data structures and creating derived data fields.
- Support for issuing queries to SQL databases and mapping query results into prefuse data structures.
- Simple, developer-friendly APIs for creating custom processing, interaction, and rendering components.

Download Link: http://prefuse.org/download/

2.5. LodLive

It is a web based application and provides visualization for linked open data. It has options to use numerous SPARQL endpoints like DBpedia and freebase. Also provides the option to use other resource address.It simply asks to select the resource and with a simple start button you can start working on the dataset.

Website Link: http://en.lodlive.it/

2.6. gFacet

Explore RDF data by combining graph-based visualization with faceted filtering techniques. The facets are represented as nodes in graph visualization and can be interactively added and removed by the users in order to produce individual search interfaces. Even multiple and distantly connected facets can be integrated in the graph facilitating the access of information from different user-defined perspectives. gFacet is based on the open source framework Adobe Flex and uses SPARQL queries to access RDF datasets.

Download Link: https://code.google.com/p/gfacet/

2.7. RDF:SynopsViz

RDF:SynopsViz is a tool for hierarchical charting and visual exploration of RDF & Linked Open Data. Hierarchical RDF exploration is based on the creation of multiple levels of hierarchically related groups of resources based on the values of one or more properties. The adopted hierarchical model provides effective information abstraction and summarization. Also, it allows efficient -on the fly- statistic computations, using aggregations over the hierarchy levels.

Website Link: http://83.212.97.83:8084/

2.8. Sgvizler

Sgvizler is a JavaScript which renders the result of SPARQL SELECT queries into charts or html elements. The name and tool relies on and/or is inspired by SPARQL, Google Visualization API, SPARQLer, Snorql and Spark. All the major chart types offered by the Google Visualization API are supported by Sgvizler. The user inputs a SPARQL query which is sent to a designated SPARQL endpoint. The endpoint must

return the results back the query results in SPARQL Query Results XML Format or SPARQL Query Results in JSON format. Sgvizler parses the results into the JSON format that Google prefers and displays the chart using the Google Visualization API or a custom-made visualization or formatting function.

PS: Queries for visualization results have to be written manually.

Website Link: http://dev.data2000.no/sgvizler/

2.9. Visualbox

Visualbox is a "special edition" of LODSPeaKr focused in allowing people to create visualizations using Linked Open Data.

PS: Requires configuring a VM.

Website Link: http://alangrafu.github.io/visualbox/