D2V: A Tool for Defining, Detecting and Visualizing Changes on the Data Web

Overview

The dynamic nature of data on the Web gives rise to the need of understanding and analyzing the dynamics of individual evolving datasets. As a matter of fact, the value of a dynamic dataset lies not only in its content, but also in its evolution history, which in some applications (e.g., trend analysis and identification), may be more important than the data itself. **D2V** is a research prototype for **detecting, analyzing and visualizing the dynamics of Linked Open Data (LOD)**, and has been developed and used in the context of the EU projects DIACHRON and IdeaGarden to **study the evolution of various datasets**, mainly from the biomedical domain (Experimental Factor Ontology - EFO, Gene Ontology - GO, and others).

The change detection functionality of D2V allows remote users and curators to identify changes, even if they have no access to the actual change process. D2V also empowers users to perform sophisticated analysis and visualizations on the evolution data, so as to understand how datasets (or parts of them) evolve, and how this evolution is related to the data itself. Our tool aims to become a critical addition to the arsenal of data analysts and scientists for dynamicity analysis in biomedical or other datasets.

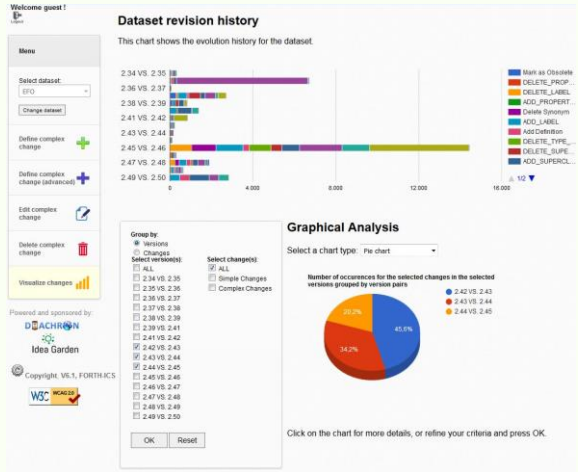

Fig.1 Analyzing the evolution of the EFO ontology

Fig.2 Analyzing the evolution of a term

Target Domains

D2V is applicable over any evolving ontology. As a result, it is **suitable for any application domain that involves ontologies that dynamically evolve over time**, even though it has been currently applied mainly in biomedical ontologies. The system is addressed to the curators of an evolving ontology, who want to understand how their ontology (or parts thereof) evolves by providing them with analytics and visualizations.

Description

D2V detects the changes between different versions of an ontology using simple SPARQL queries, and stores these changes in an appropriate structure. Then, the detected changes are presented to the user through interactive interfaces and visualization paradigms (see Figure 1 for a visualization of EFO evolution). We provide different views of the evolution history for different types of analyses: the user can see the evolution history of a given URI (**term-centric view** – e.g., return all changes associated with a specific URI in the ontology, see Figure 2), of the dataset as a whole (**dataset-centric view** – e.g., view all changes in the Disease Ontology), or of specific versions (**version-centric view** – e.g., view all changes between a given pair of versions); or the user may be interested in a **change-centric view**, where the instantiations of a given change are reported (e.g., return all classes that were made obsolete); the user can filter the different results to a fixed set of changes or versions; or visualize evolution along a series of consecutive versions, or for an arbitrary pair only.

To further help the user in analyzing the dataset evolution, the system allows the definition of custom changes for a more accurate analysis. In particular, D2V handles **two types of changes**, with the aim of making changes intuitive and human-understandable. The first is **simple changes**, which are fine-grained changes defined at design time that provide formal guarantees on the soundness and completeness of the detection process. The second type is **complex changes**, which are custom and defined at run-time by the user to satisfy application-specific needs; for example, complex changes may be used to report coarse-grained changes, changes that are important for the specific application or user, changes with special semantics, or changes that should not happen at all (their detection being like an alert for an abnormal situation). D2V provides also functionalities for creating, editing or deleting complex changes.

Additional Information

A **video** demonstrating D2V in action is available at:
https://www.youtube.com/watch?v=oY7qBBfcHYg

A **demo** of D2V is available at:
http://139.91.183.40:8080/D2VSystem

More details can be found at the following publications:

[1] Y. Roussakis, I. Chrysakis, K. Stefanidis, G. Flouris, Y. Stavrakas. A Flexible Framework for Understanding the Dynamics of Evolving RDF Datasets. Best Student Paper Award. ISWC-15, Research Track.

[2] Y. Roussakis, I. Chrysakis, K. Stefanidis, G. Flouris. D2V: A Tool for Defining, Detecting and Visualizing Changes on the Data Web. ISWC-15, Demonstrations Track.

D2V video

Contact details: *www.ics.forth.gr/isl*

Ioannis Chrysakis **hrysakis@ics.forth.gr**

Yannis Roussakis **rousakis@ics.forth.gr**

Giorgos Flouris **fgeo@ics.forth.gr**