Visualizing Research Networks' Evolution over Time

Anastasia Dimou, Laurens De Vocht,
Ruben Verborgh, Erik Mannens, and Rik Van de Walle

Ghent University - iMinds, Belgium {firstname.lastname}@ugent.be

Abstract. Linked Data visualizations present a rather static view of entities and their relations, despite the semantically rich data. In contrast to the spatial dimension, which is well-exploited, the temporal dimension is neglected. In this demo, we address the temporal dimension, enabling time-dependant visualizations, applied to research networks based on semantically annotated research metadata. Such visualizations can reveal insightful information otherwise not able to be observed.

1 Introduction

Non-Semantic Web experts lack in-depth knowledge of Semantic Web technologies and Linked Data: they cannot form queries, nor traverse the semantically annotated information. Thus, they cannot conceive the potential this interlinked information offers. Linked Data visualizations aim to support those users, yet most visualizations offer a rather *static view* of entities and their relations. Even though the Linked Data spatial dimension is well-exploited so far, e.g., visualizing Linked Data on different geographical maps, the temporal dimension is neglected. Most visualizations focus on a snapshot, i.e., the latest state of all data accumulated, or just list time-stamped data on timelines or calendars.

With this demo, we propose interactive visualizations which take into consideration the Linked Data temporal dimension. In our exemplary case, visualizations reveal the existence and the nature of research networks, as well as the collaborations and the emerging communities of practice, as they evolve over time. They thereby provide insightful information regarding the evolution of patterns and trends over time, that allow non-Semantic Web experts to identify relationships and draw conclusions that could not be otherwise observed.

2 State of the Art

In most cases, visualizations are presented in a complementary fashion to the information presented. In terms of graph-based research network visualizations, the Semantic Web Journal [6], for instance, relying on its own dataset, provides research network visualizations as researchers author papers together or review

different submissions. ArnetMiner [9] distinguishes the networks (unlinked co-authors) from the communities of researchers (linked co-authors) into two different types of visualizations. ResXplorer [4] differs significantly, as its visualizations provide the optimal path between two entities (researchers, publications, papers, or conferences) presenting the same graph from different perspectives. However, none of these solutions capture the research networks evolution over time as our visualizations do, nor provide alternative paths between two specific researchers.

Shneiderman [7] proposed seven basic data-based types of visualisations, assuming that users view collections of items, where items have multiple attributes. Temporal datatype is listed among them, covering (i) timestamped data (i.e., a *precise moment in time*), but also (ii) items *spanning through time* with defined starting and end date. Proportionately, temporal data is suitable to be displayed in *calendars* and *timelines*, but time-series data is a perfect candidate to be *mixed with other dimensions*, allowing new analyses over data changing over time to *detect patterns and trends*. Most existing solutions focus on the calendar/timeline type of visualizations. For instance, the Linked Data Visualization Model (LDVM) [2] mapped temporal datatypes to timelines and calendars, in the same way as LDVizWiz [1] and Sgvizler [8] that visualize data's temporal dimensions only with timelines, and VizBoard [10] that visualize them in calendars. On the contrary, our visualization suite provides *both* types of visualizations. In this demo, we focus on time as a dimension to graph visualizations, showcasing them with research networks, allowing to observe changes as time passes.

3   Research Network Visualization
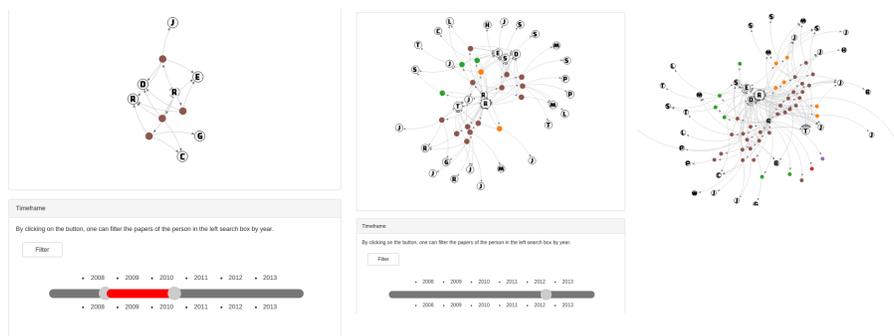
Fig. 1. The research network of an actual Semantic Web researcher:
(a) *time span:* from 2008 to 2010, (b) *precise time:* in 2012, and (c) *overall*

Our visualizations suite, LODVizSuite [5], supports two types of research network visualizations: (i) **weighted graph-based with center** displaying the research network around a researcher, **in the dimension of time or not**, and (ii) **path-based** displaying alternative paths between two particular re-

searchers. A live demo of the research networks visualizations is available at
`http://ewi.mmlab.be/time/`.

The first type of research network visualizations relies on the researcher's
publications and co-authors. The network is represented as weighted graph with
center the researcher whose research network is examined. For instance, the
overall network of collaboration of an actual Semantic Web researcher appears
in (c) of Figure 1 based on accumulated data without taking into consideration
the time dimension. A timeline is introduced to adjust the time dimension. The
research networks can then display the network (i) at a **precise time**, e.g.
in 2012, or (ii) in a **time span**, e.g., between 2008 and 2010, (a) and (b) of
Figure 1 respectively. Relying on different time snapshots, it is easy to observe
the collaborations evolution over time. For instance, how the network grows
over time, how it is shaped in a certain time-span, which collaborations are
reinforced and which disappear. While the first type of visualizations is focused
on a single researcher and eventually established collaborations, the second type
is focused on the network formed between two researchers, displaying alternative
opportunities to establish collaborations. The visualizations display alternative
paths of collaborations between them, considering series of common publications
and co-authors as intermediate steps. The visualizations, in this case, rely on an
optimized path-finding algorithm, the Everything Is Connected Engine [3].

We implemented visualizations of the temporal dimension to our LODVizSuite
suite. We demonstrate it with the RILOD dataset.

Dataset. The visualizations were applied on top of the Research Information
Linked Open Data (RILOD) dataset[1] that contains information about 22,006
researchers, 2,022 research groups, 24,635 projects and 608,729 publications,
from 2008 till 2013, resulting in more than 400,000,000 triples.

Vocabularies. The data is annotated with the Common European Research Infor-
mation Format (CERIF)[2] and the EWI Linked Open Data Ontology (EWILOD)[3].

Implementation. The visualizations are part of the LODVizSuite, a suite of data
and schema-agnostic visualizations, easily applicable and transferable across dif-
ferent datasets, offering an interactive interface on top of Linked Data. At client
side, the LODVizSuite consists of a number of JavaScript libraries relying on
D3.js[4] library, to display different types of visualizations. At server side, a Java-
based implementation handles and processes the client's requests and retrieves
in real time the corresponding Linked Data according to SPARQL queries pub-
lished at a dedicated DataTank[5] instance. The DataTank publishes data on the
Web by transforming them in an HTTP API, among them SPARQL templates. Pa-
rameters can be passed to the SPARQL template at request time, which replace
placeholders and constructs a valid SPARQL query. The visualizations are easily
applicable to other datasets, only by adjusting the published SPARQL queries.

---

[1] `http://goo.gl/zKZbBM`
[2] `http://www.eurocris.org/ontologies/cerif/`
[3] `http://ewilod.be/ewilod/lod/0.1/ontology`
[4] `http://d3js.org/`
[5] `http://TheDataTank.com`

Usage. A user can access the visualizations and start typing a name in the search box. While typing, an auto-completion menu appears and the user may select the entity, in this case a researcher, to obtain the corresponding visualization. Then using the time frame at the bottom, the user may obtain a time snapshot by select either a precise time or a time span. Firstly, the user needs to indicate the end time by clicking on one of the available dates, appearing under the time slider. Then, he may select the start date, by clicking on one of the available dates, appearing on top of the time slider. If the start and end date are the same, a visualization for this precise time is shown. If they are different, a visualization of the corresponding time span is shown.

In this paper, we demo how, taking advantage of the Linked Data temporal dimension captured, we can enable time-dependent visualizations and let users identify possible patterns and trends, otherwise not able to be observed.

References

1. G. A. Atemezing and R. Troncy. Towards a linked-data based visualization wizard. In *ISWC 2014, 5th International Workshop on Consuming Linked Data*, 2014.
2. J. M. Brunetti, S. Auer, and R. García. The linked data visualization model. In *International Semantic Web Conference (Posters & Demos)*, 2012.
3. L. De Vocht, S. Coppens, R. Verborgh, M. Vander Sande, E. Mannens, and R. Van de Walle. Discovering meaningful connections between resources in the Web of Data. In *Proceedings of the 6th Workshop on Linked Data on the Web*. CEUR-WS, 2013.
4. L. De Vocht, E. Mannens, R. Van de Walle, S. Softic, and M. Ebner. A search interface for researchers to explore affinities in a linked data knowledge base. In *Proceedings of the 12th International Semantic Web Conference Posters and Demonstrations Track*. CEUR-WS, 2013.
5. A. Dimou, L. De Vocht, G. Van Grootel, L. Van Campe, J. Latour, E. Mannens, P. Mechant, and R. Van de Walle. Visualizing the information of a Linked Open Data enabled Research Information System. In *Proceedings of the Current Research Information Systems Conference*, 2014.
6. Y. Hu, K. Janowicz, G. McKenzie, K. Sengupta, and P. Hitzler. A Linked-Data-Driven and Semantically-Enabled Journal Portal for Scientometrics. In *Proceedings of the 12th International Semantic Web Conference*. 2013.
7. B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Visual Languages. Proceedings., IEEE Symposium on*, 1996.
8. M. G. Skjæveland. Sgvizler: A JavaScript Wrapper for Easy Visualization of SPARQL Result Sets. In *9th Extended Semantic Web Conference (ESWC 2012), workshop and demo proceedings.*, 2012.
9. J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. ArnetMiner: Extraction and Mining of Academic Social Networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2008.
10. M. Voigt, S. Pietschmann, and K. Meißner. A Semantics-Based, End-User-Centered Information Visualization Process for Semantic Web Data. In *Semantic Models for Adaptive Interactive Systems*. 2013.