SemaZoom: Semantics Exploration by Using a Layer-Based Focus and Context Metaphor

Dirk Burkhardt, Kawa Nazemi, Matthias Breyer, Christian Stab, and Arjan Kuijper

Fraunhofer Institute for Computer Graphics Research, Fraunhoferstraße 5,
64283 Darmstadt, Germany
{dirk.burkhardt,kawa.nazemi,matthias.breyer,christian.stab,
arjan.kuijper}@igd.fraunhofer.de

Abstract. The Semantic Web is a powerful technology for organizing the data in our information based society. The collection and organization of information is an important step for showing important information to interested people. But the usage of such semantic-based data sources depends on effective and efficient information visualizations. Currently different kinds of visualizations in general and visualization metaphors do exist. Many of them are also applied for semantic data source, but often they are designed for semantic web experts and neglecting the normal user and his perception of an easy useable visualization. This kind of user needs less information, but rather a reduced qualitative view on the data. These two aspects of large amount of existing data and one for normal users easy to understand visualization is often not reconcilable. In this paper we create a concept for a visualization to show a bigger set of information to such normal users without overstraining them, because of layer-based data visualization, next to an integration of a Focus and Context metaphor.

Keywords: Human-Computer-Interfaces, Semantics Visualization, Information Visualization, Semantic Web.

1 Introduction

With the development of the Semantic Web huge knowledge bases were created to use the features of storing data in a structured and linked form. Typically a knowledge base consists of an ontology in the format of RDF or OWL. Nowadays adequate visualizations are available for visualizing the schema and its instances. The most common visualization type for semantic data is a graph-based visualization, because it is one of the most known and understandable visualization metaphor for normal users. An actual and important challenge is the visualization of massive data in adequate form, because today knowledge bases containing more than 100.000 instances and sometimes also more than 2 million instances. Of course, the visualization of all the available instances is not appropriate, because it will overstrain the user by visualizing too many entities. But by exploring and searching within huge knowledge bases the resulting number of instances is sometimes more than 50 instances and it is hard to automatically reduce it to a lower value. So in these cases a visualization or a visualization metaphor is needed for visualizing a higher numbers of entities.

In the research area of information visualization some approaches exists for providing the possibilities to present also higher numbers of entities in an acceptable way. Especially interaction metaphors like Focus and Context are established forms for visualization large numbers of nodes. A typical and easy to use representative is the Fisheye View, where during the interaction time the entire context is visible, next to a parallel zoomed area in which a detail view on the data is given. The interaction metaphor has as central approach the use of a magnifier, that zooms the area of interest e.g. around the mouse. The more an entity is away from the area of interest, the smaller it is displayed, so for instance entities on the borders of the visualizations are very small. Over all the time, the entire context of the presented data is visible. So also the relations between the elements are visible all the time. With such an interaction metaphor, the user is able to explore also larger subsets of a dataset by regarding the context of the entities.

In this paper we are going to present a concept of a Focus and Context visualization named *SemaZoom*. With *SemaZoom* a graph of instances from an underlying ontology will be visualized and can be explored. It uses the Fisheye Effect for providing the possibility to read all entities and regarding that the context is visible all the time. *SemaZoom* contains also another new approach, for presenting concepts and instances in a different way. Normally a semantic structure is visualized in a single graph, presents the concepts and instances equally. This approach has as disadvantage that many nodes (including the concepts) will be presented, but often only the instances are relevant. So the user sees also objects that are not relevant for him. In fact there are also graphs, which only present instances. But the corresponding concept often contains additional and relevant information, which can be important for the user. For instance, if a user searches for golf, the concept can indicate if the instance will be the sport activity or the automotive. For this reason we extended the graph by using fisheye effect with a second layer that will show the concept to a presented instance. So the *SemaZoom* presents the semantic data from an ontology in a graph of instances by using the fisheye effect and in a second layer with drawn colored background shapes surround the instances from the first layer. These shapes address concrete concepts. The names of the concept can be written directly on the shapes or in a legend next to the shapes. This approach guaranties that also additional information from the schema level are available for the user, during his exploration through the instance data. Furthermore the space for presenting information will be exhaustively used in efficient way, thus reducing the cognitive overload of the user on a low level.

2   Related Works

For visualizing of large amount of data, different kinds of established visualization approaches are available. Especially by visualizing semantics data are taxonomy-based and graph-based layout techniques appropriate. Furthermore, different kinds of interaction techniques and metaphors are available for realizing the interaction and exploration through the data. One of the most trivial form of visualizing data are the presentation of all available nodes, but in particular by massive data this approach is not adequate, because the users will get overstrained. Today different approaches exist to avoid an overstraining of the users, for instance the reducing of the data on

relevant nodes or efficient visualizing strategies to use the space on a screen optimal. The approach for reducing data on relevant data is a critical once, because for this a method is required, which is able to determine relevant and irrelevant nodes and to avoid that the systems hides nodes, which are necessary for the user. So in this paper we are concerning on strategies, that will not hide nodes, but rather allow to define a focus on the data. This is only possible with zoom and filter techniques basing on the visualizing level.

One visual technique for providing a filter technique is Panning. First systems provide this feature by scrolling, panning or paging [5]. The underlying idea bases on moving a viewport over a document. The most known representative is scrolling, which is realized in application by scrollbars. In general panning can be realized with 4 different main interaction forms [6]: with scrollbars, by using a dragging metaphor, by using the map metaphor and by a combination of map and pointer.

By large numbers of items or nodes panning will be no adequate metaphor for interacting and exploring through the data, because the scrolling etc. to another position in a document takes a long time. If zoom functionality will be added, the browsing through a document or in general through the dataset is much faster, because of the smaller size of nodes that represents information over which has to be browsed [5]. For a general use, the zoom function allows the user to get an overview of the existing data and when he needs more information, he can take a "deeper look" into the data by zoom-in to the data elements of interest.

A disadvantage by providing a zooming functionality is the loose of the context, when the user zooms into a data area of interest. An approach that compensates this aspect is the use of Multiple Views. By this approach the user has at least two views on the data. On for the main interaction, often this is a bigger view area, to work within the data. Next to it, there are often one or more other views, which provide an overview to the entire data set. This metaphor is commonly used for geographical information e.g. GoogleMaps, where a user has und the bottom right side a small overview window from the map e.g. from a country or continent.

Next to the general metaphors for working with large datasets, there are also solutions which combining some of these basic metaphors. A result of such a combination are so called Focus and Context metaphors or also known as Overview and Detail metaphors. An overview to a number of established techniques for Focus and Context is given in Cockburn et al. 2008 [1].

A common known representative for Focus and Detail is the Fisheye view effect. One of the first implementation on a graph-based visualization is described in Sarkar et. al. [3]. The fisheye effect makes nodes bigger within a focused area than nodes which are out of this area. In general the nodes are becoming smaller, the more they are distanced from the focused area. In fact all elements of the graph are visible during the whole time of interaction, which regards the visibility of the context. Additionally the user is able to see and read a node from the graph, if the node is within the focus are e.g. by moving the mouse over it.

An example implementation from fisheye view on graph-based visualization for showing semantics data is presented in Zhang et. al. [2]. In this implementation the schemas and instances are handled equally, so that the benefit of a structured information source will not be full supported, but with this kind of implementation the user gets an easy form to interact through a bigger network of semantically information.

3   Concept of the SemaZoom Visualization

Graph-based visualizations are most common used visualization type for visualizing semantics data. They provide an easy view on the data and furthermore they allow an easy and understandable navigation through the concepts and instances. But semantics data-sources are often very voluminous so that a useful navigation through this massive data is impossible.

This aspect is the challenge, which the *SemaZoom* navigations try to solve, next to an effective presentation of the different objects that exist in semantic data structures for concepts and instances.

3.1   Differentiation between Concepts and Instances

In the scenario of an information search, the required results are often concrete instances, which describing explicitly specific information. Concepts are commonly used for filtering instances from a given result set. The reason is that concepts in the context of semantic structures defining general characteristics for concrete occurrences in form of instances, which can be classified by properties, but of course the values can be differently for each instance. In our visualization framework we are using for filtering the result set by selecting or deselecting concepts other kinds of visualization like the *SeMap* [7]. As a consequence for supporting the simultaneous use of visualization for providing different possibilities and regarding different aspects, with our SemaVis-framework it is possible to composite a personal so called "Knowledge-Cockpit" [4] by coupling preferred visualization (see also the screenshot in Fig. 3.).

The focus of this visualization lays on the visualization of the instances and its relations. Because of the fact that instances are linked with each by semantically relations, graph-layout algorithms are suitable. In comparison to graph-based visualization, which are showing concepts and instances in a same way (e.g. ZoomRDF [2]), we only visualize instances in the graph, which allows us to present more instances without overstraining the user. This graph consisting of semantically instances and relation is presented in one layer.

An assignment of instances to their concept provides an orientation for the user, because the concepts implicitly containing additional information to an instance. This additional information to an instance is not required during most of the time by interacting or exploring through the data, so that the concepts do not have to be presented conspicuously. For achieving such a restrained presentation, we show conceptual information like the information of altitudes values are presented on maps. We realize it by coloring the background in dependence of the assigned concept. This coloring will be done in a second visual layer, which is underneath the Instance and Relation Layer (see Fig. 1.). So the background of every instance is colored, by an existing legend it is possible to identify the concrete concept, where an instance is assigned to. Semantically concept relations cannot be visualized with this kind of visualization. The semantically concepts in form of background shapes are drawn into the second layer, which is placed under the instance layer. Shapes will be updated or redrawn automatically, if the positions of the instances are changing.
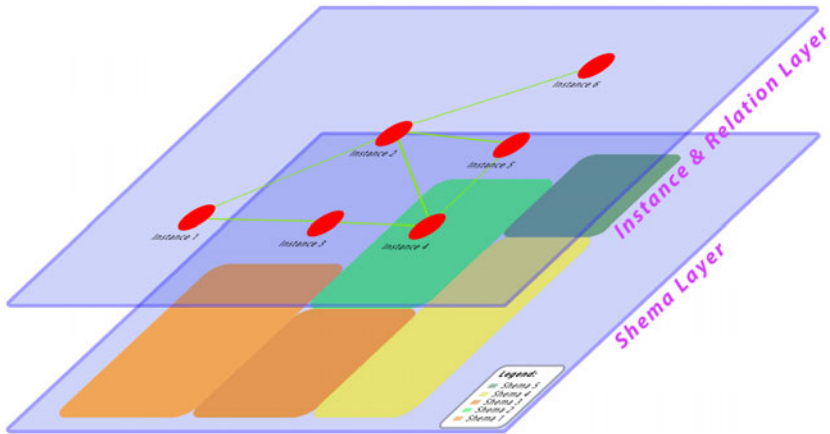
Fig. 1. Visual concept for *SemaZoom* to visualize semantic instances and concepts

3.2   Usage of Focus and Context Metaphor

By the presentation of semantics data, e.g. from an ontology management system, are often very large data sets are to visualize. For supporting an adequate visual presentation of such large data sets within graph-based visualization some implementations do exist by using a Focus and Context view. A well-known and with often positive results evaluated metaphor is the Fisheye-view metaphor.

We use the fisheye effect within the *SemaZoom* visualization for providing an interactive data presentation. So the user can navigate through big dataset to search explore for information. Because of the use of the Fisheye view, the user is able to see focused information in a normal size so that they are good readable, but he also has permanently an overview about the context, how an instance is related to another once. Because of the changing size and position of every instance by the use of such a dynamic zoom-metaphor, the background shapes that represents the concepts, have dynamically be updated, to ensure that these information will be shown correctly during the entire time of interaction.

For the generation of the fisheye view effect, we are using a similar calculation, like it is used in Sarkar et. al. [3], we only changed some parameters, which are providing a better interactivity for a web-based usage.

3.3   Implementation and Feature Summary

As a result we implemented the concept prototypically as a visualization module that can be used in our developed SemaVis-framework. An impression of the implemented visualization is given in Fig. 2.

This kind of visualization supports the presentation of instances, instance relations and concepts as semantic objects. For the navigation the visualization supports multiple paths, where a user does not have to go back for changing to another instance. In the current version *SemaZoom* is only less configurable, which makes it not recommendable for the use in adaptive systems. Actually only the color of the background shapes, which represents the concepts, and the color of the instance relations can be changed. A summary to the supported features is given in Table 1.
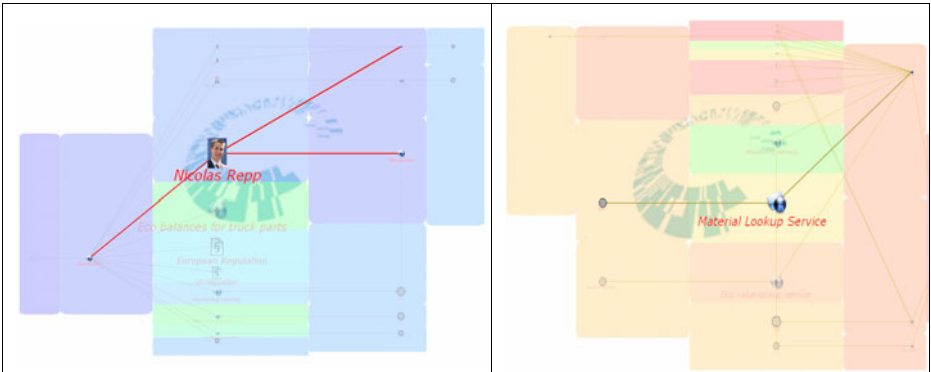
Fig. 2. Screenshots of the *SemaZoom* with different configurations

Table 1. Summary of supported semantics data features by this type of visualization

	Semantics					Navigation		Adaptation			
	Concepts	Individuals	Relations	Multiple Inheritance	Properties	Single path	Multiple path	Size	Color	Order	Predictions
SemaZo	✔	✔	✔	-	-	✔	✔	-	(✔)	-	-

4   Case Study

The described *SemaZoom* visualization was developed as a part of the Core-Technology-Cluster (CTC) of the THESEUS Program (Theseus 09), a 60-month program partially funded by the German Federal Ministry of Economics and Technology. The partners in the THESEUS Program research under the device "New Technologies for the Internet of Services" heterogeneous technologies for gathering and offering semantic information on web. The program itself consists of twelve projects, divided in THESEUS Use Cases and the THESEUS Core-Technology-Clusters. Where the six Core-Technology-Clusters are led by research institutions and focus on fundamental research areas, the THESEUS Use Cases are led by enterprise institutions and bridge the gap between fundamental research and industrial dissemination. Different enterprise partners focus on their usage scenario of the different areas of information processing. For example the Siemens Corporation investigates the processing of medical-related information. In this THESEUS Use Case (Medico) different usage scenarios identify different user groups: There are medical doctors, who use the information of the patient's clinical and medical records to find similar cases and provide the adequate care for them. On the other hand you have the patients themselves, who should be able to understand about their disease and find for example groups or a community with similar ailments.

Another example for a THESEUS Use Case is THESEUS Texo, which is conducted by SAP Research Germany and investigates different models and techniques for providing, engineering, disseminating and using web-based services. The different user roles e.g. service-engineers, domain-experts or service-consumers handle the same information in different ways. Complex information structures defined as ontologies makes it necessary to provide here a best fit of visualizing the information structures for the different user-groups and their precognitions.

A third example for a THESEUS Use Case is Contentus, led by the German National Library. Here you find the same heterogeneity of users. There are domain-experts, who have the required knowledge in a specific scientific domain, e.g. experts for German Literature, but are not experts in using and processing complex ontology-based information-systems. Of Course you find in Contentus the average user too, who just explores knowledge domains and expects a very simple to use visualization and user interface.

Beyond the THESEUS Use Cases there are six THESEUS Core-Technology-Clusters (CTCs) investigating different fundamental research questions regarding semantic information processing. The CTC are mainly conducted by research institutes. The CTC for Ontology Management, led by "Forschungszentrum Informatik", investigates for example managing, reasoning, editing and inferencing ontologies. The CTC Situation Aware Dialog Shell investigates different questions regarding context-aware information processing.
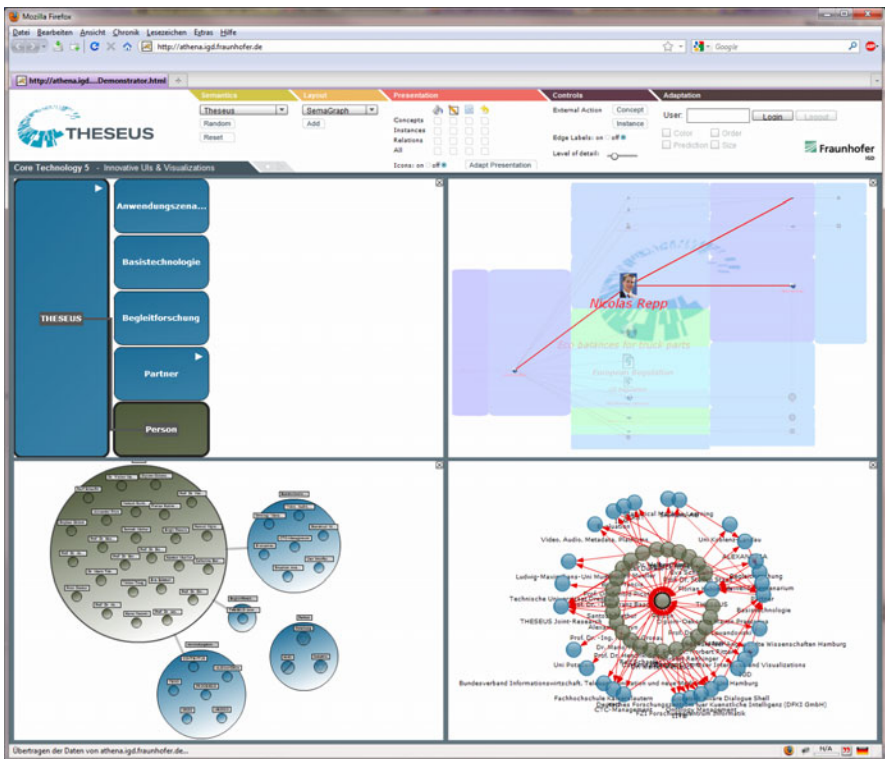

Fig. 3. *SemaZoom* integrated into the SemaVis Knowledge-Cockpit

The *SemaZoom* visualization was developed as a CTC-component of the THE-SEUS Core-Technology-Cluster Innovative User Interfaces and Visualizations and is used in different THESEUS Use Cases for exploring knowledge domains within the specific scenarios of the Use Cases. *SemaZoom* was tested in different Use Cases in their specific semantic knowledge domain. The first tests in real scenarios, where users were able visualize medical documents and information in the UseCase Medico. Because of the high number of related items, with *SemaZoom* it was possible to explore coherence documents and reports to a patient. We also integrated the *SemaZoom* visualization into our demonstrator with which a user can generate is own personal "Knowledge Cockpit" (see Fig. 3.).

5   Discussion

Interacting through semantics data is an up growing challenge, because the currently existing approaches and implementations are designed for specific aspects, often they are focusing on technical details, for instance to highlight ontology-based features for organizing the data in an ontology management system. Normal users, respectively experts of foreign domains, need other kinds of visualizations. With this *SemaZoom* visualization we tried to provide an easy to user interface for semantics data, but during the conceptualization and also during the implementation we get some challenges, where further work is required and useful. A general aspect is the classifications of the user profiles, particularly depending to their IT skills. Depending to those user profiles a mapping is necessary, which kinds of visualization will be "easy to use". A very easy visualization metaphor can be a helpful tool for computer novices, but for advanced users or experts it can be an annoying tool, because it provides to less information.

This is also a necessary question for the *SemaZoom* to define which information is useful and should be presented to a user during his interaction with the visualization. Currently we provide information only to the name of an instance or concept, but no further details to or surround this object. Also the names of the relations between instances are currently hidden, because it can be confusing if in general too many text labels will be presented, especially if a large data set is to visualize.

A specific challenge for this concrete visualization is the form, how the semantic schema with its hierarchy of concepts can be presented and also their relation to each other. In this version we show only the concrete concept to an instance, but if another instance from a higher or lower concept level is presented next to it, it can be confusing for the user. An approach to avoid this can be the use of different brightness levels as an indicator for the level within the schema hierarchy, but for this the entire schema must be known by the visualization, which generates a high overhead especially by large semantics databases, because all the information must be known by the visualization. This circumstance makes it hard for the use of client-server based search technologies.

6   Conclusion

In this paper we presented a concept for a layer-based Focus and Context metaphor for the usage on semantics data and for supporting a semantics exploration. The concept consists of two steps. First, we generate the presentation of the data with a

graph-based visualization that shows the instances and instance relations in one layer and concepts in a second layer. The concepts contain additional information to the presented instances, like the type of it e.g. an automotive, a sport activity etc.

We also integrated a Focus and Context metaphor, which allows the presentation of a large number of instances without overstraining the user, because these kinds of metaphors provide the feature to show an overview of the data, next to details to viewed nodes of interests. In our visualization we are using the Fisheye View metaphor as one representative of Focus and Context. As a result of this Fisheye View effect, the user can explore parts of the visualized data by focusing it and the elements within the focus area will be shown in a normal well-readable size. Objects out of the focus area are presented with a very small size, but they are visible all the time, so that the context – especially the instance relations - is always observable for a user.

By combining both aspects, we created a new form for a semantics visualization with an interactive style. The *SemaZoom* visualization provides an easy understandable presentation of semantics, which also allows the possibility to explore bigger datasets. In particular by the visualization of semantics data, this kind of visualization provides a well overview on the data and parallel the user is able to look in detail on instances of interests.

References

1. Cockburn, A., Karlson, A., Bederson, B.B.: A review of overview+detail, zooming, and focus+context interfaces. ACM Comput. Surv. 41(1), Article 2 (2009)
2. Zhang, K., Wang, H., Tran, D.T., Yu, Y.: ZoomRDF: semantic fisheye zooming on RDF data. In: Proceedings of the 19th International Conference on World Wide Web (WWW 2010), pp. 1329–1332. ACM, New York (2010)
3. Sarkar, M., Brown, M.H.: Graphical fisheye views of graphs. In: Bauersfeld, P., Bennett, J., Lynch, G. (eds.) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1992), pp. 83–91. ACM, New York (1992)
4. Nazemi, K., Burkhardt, D., Breyer, M., Stab, C., Fellner, D.W.: Semantic Visualization Cockpit: Adaptable Composition of Semantics-Visualization Techniques for Knowledge-Exploration. In: International Association of Online Engineering (IAOE): International Conference Interactive Computer Aided Learning 2010 (ICL 2010), pp. 163–173. University Press, Kassel (2010)
5. Gutwin, C., Fedak, C.: Interacting with big interfaces on small screens: a comparison of fisheye, zoom, and panning techniques. In: Proceedings of Graphics Interface 2004. ACM International Conference Proceeding Series, vol. 62, pp. 145–152. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario (2004)
6. Kaptelinin, V.: A comparison of four navigation techniques in a 2D browsing task. In: Conference Companion on Human Factors in Computing Systems, CHI 1995, pp. 282–283. ACM, New York (1995)
7. Nazemi, K., Breyer, M., Hornung, C.: SeMap: A Concept for the Visualization of Semantics as Maps. In: Proceedings of the 5th International Conference on Universal Access in Human-Computer Interaction (HCII 2009), pp. 83–91. Springer, Heidelberg (2009)