The Role of Semantic Data in Engineering Interactive Systems

Jürgen Ziegler and Timo Stegemann
University of Duisburg-Essen, Interactive Systems Research Group

47048 Duisburg, Germany
{juergen.ziegler, timo.stegemann}@uni-due.de

ABSTRACT

Semantic technologies, as advocated and used in the context of the Semantic Web, are increasingly seen not only as a pre-requisite for the automated processing of distributed data on the Web, but also as a basis for new methods and tools for engineering interactive systems. In this position paper, we briefly review the state of the art related to visualizing and exploring semantic data, in particular Linked Open Data, as well as techniques for developing user interfaces for semantic data. We also refer to some of our own work in this field. Based on this overview, we present a range of open research issues that we organize around three paradigmatic concepts that might inform and guide future developments in this area. The three paradigms discussed in this paper are: constructive exploration, dynamic contextualization, and transparent integration of semantic data. These conceptual contributions together with the concrete developments described shall serve as input to a broader discussion of a roadmap for the future engineering of interactive systems.

Keywords

Semantic Web, visual exploration, context-adaptation, semantic widgets

INTRODUCTION

Since the first articulation of the vision of a Semantic Web (Berners-Lee et al., 2001), the amount of data on the Web represented with semantic techniques has increased enormously. In contrast to the conventional Web which links web pages as complete documents, the Semantic Web links data based on a uniform data model. This model consists of elementary statements in 'subject-predicate-object' form, expressed with RDF, and uses shared vocabularies defined as ontologies by means of RDFS and OWL. Linking the individual statements results in single, huge data graph, which is often referred to as the Web of Data.

Today, large amounts of semantic data are available openly on the Web, often automatically extracted from conventional information sources such as encyclopedias (e.g. DBPedia which is extracted from Wikipedia, geographical databases, bibliographies, product data and many others (for an overview see, e.g. Bizer et al.,2008). The LOD (Linking Open Data) initiative has invested considerable effort in interconnecting different datasets, resulting in the LOD Cloud which has meanwhile reached an enormous size. Beyond general public information sources, semantic techniques are also making their way into more specific application domains such as electronic commerce (e. g., Hepp, 2008) or corporate document management (e.g., in the current Microsoft Sharepoint).

While the original objectives of semantic techniques were mainly directed at making Web information machine-processable, it is increasingly recognized that linked open data and other semantic data pools are valuable information sources that can be used interactively by end users. This creates a need for methods and tools that allow users to interact with Semantic Web data directly, rather than through a web application that integrates and delivers data in standard web pages.

In contrast to conventional Web front ends, user interfaces and search tools for Semantic Web data can exploit the semantics of the underlying data structure, for instance, to let users explore the data under different perspectives and formulate more targeted and complex queries. The potential of interactively exploring semantic data has been demonstrated by a variety of tools that use the different semantic relations for browsing the data or for creating search facets that allow users to flexibly filter the data. Visual techniques for formulating complex queries have been developed, for instance, in prior work of ours (Heim, Ertl, & Ziegler, 2010; Heim et al., 2009). In addition to providing improved search and exploration capabilities, however, semantic representations can also enable users to construct their own views of a semantic data pool, which can either be used for exploring the data in some user-defined visual configuration or for presenting (and potentially editing) a more or less complex cut-out of the large RDF based data graph.

Semantic techniques open up new approaches to end-user tailoring or even development of interactive applications. Due to the explicit and user-inspectable representation of semantic models in the form of ontologies, to which the (instance) data are directly linked, a range of new development methods and tools can be conceived which also have implications for the architecture of interactive systems. In this position paper, we describe some of the new conceptual aspects of engineering interactive systems based on semantic data which may change the way interactive systems and user interfaces will be developed in the future. In the following, we first present a brief review of the state of the art and describe some of our own developments. We then discuss three lines of research in which we believe semantic techniques will have an impact on engineering interactive systems in the future.

RELATED WORK

A variety of tools have been developed in recent years that allow users to browse Semantic Web data directly. Well-known examples of Semantic Web Browsers include Tabulator (Berners-Lee et al., 2008), Parallax (Huynh & Karger, 2009), and Humboldt (Kobilarov & Dickinson, 2008). Different models and metaphors have been proposed which can roughly be subdivided into page-based and graph-based techniques. Page-based approaches show a resource as a single Web page with links representing the relations to other resources while graph-based approaches provide a visualization of the RDF graph directly. RDF data also lend themselves well to faceted browsing since related concepts can be used to filter the instance set of some target concept. This technique has been used in a number of systems in which facets were mainly represented as menus or links on pages showing the filtered results (as an example, see schraefel et al, 2005).

In our own work, we have developed several tools for visualizing and exploring RDF data. *Facet Graphs* combine the benefits of graph-based RDF visualizations with facetted browsing and allow users to formulate complex facetted queries in a visual form (Heim et al. 2010). In this tool, a graph visualization was extended to contain set-valued nodes, showing all instances belonging to the concept represented by the node. This allows users to formulate complex queries by selecting filter attributes from adjacent as well as distant concepts. *RelFinder* is a visual tool for detecting so far unknown relationship or patterns. RelFinder (Heim et al. 2009) is a browser-based system that can take two or more elements of an RDF dataset as starting points for extracting the relations among those elements and visualizing them as a network that can further be manipulated and explored. A marketing manager, for example, could thus identify products of competitors in a certain product category, find out whether they were used for similar purposes, or check whether they use the same suppliers. RelFinder can therefore be regarded as an interactive knowledge discovery tool.

While tools for browsing semantic data are typically generic and support the unconstrained exploration of the data, in many cases more domain- and application-specific user interfaces are needed. Several techniques have been proposed to (rapidly) develop UIs for semantic data, mostly with a focus on presentation.

The Xenon project (Quan & Karger, 2005) is an example of an approach for transforming RDF data into presentations based on a stylesheets. Based on the concept of XML-Stylesheets (XSLT), the authors developed a RDF-based stylesheet language, which defines concepts for transforming RDF-data using lenses and views. The purpose of lenses is data selection from instances whereas views describe the visualization of data elements. Just like with XSLT it is possible to embed HTML markup directly into the stylesheet for generating HTML representations of the data. As a follow-up to Xenon, Fresnel (Pietriga et al., 2006) still uses lenses for selecting data but replaces the view concept by formats that define, for instance, whether the data selected should be presented as a link, an image or as text.

OWL-PL (Brophy, 2010), is a further language for transforming RDF/OWL data into (X)HTML. The language is strongly inspired by XLST and has the main goal to provide a simple transformation language for semantic data. OWL-PL allows the combination of transformational and representational markup. The language defines stylesheets, which are connected to semantic data by using a stylesheet ontology. The ontology describes how specific RDF classes are related to stylesheet elements. With the introduction of LESS [ADD10], a complete workflow from creating and processing templates for semantic data up to sharing templates between users is described. The declarative template language LeTL (LESS template language) is specified as a Smarty based templating language. It can process and transform semantic data from RDF documents or data requested by SPARQL queries.

Work in our group has also addressed techniques for facilitating the construction of frontends for semantic data. In (Stegemann et al., 2012), we describe X3S, a technique for composing presentations of semantic data that can be created by simple drag-and-drop actions. A recent development are semantic widgets which will be further described in the following section.

RESEARCH CHALLENGES

In the following, we describe selected areas which pose particular challenges for future research related to semantic data and engineering user interfaces. To characterize the challenges involved and to indicate possible directions, we group the issues around three principles which we also see as conceptual contributions to discussing future research directions: *Constructive Exploration*, *Dynamic Contextualization* and *Transparent Integration of Semantic Data* which will be discussed in the following sections.

Constructive Exploration

Semantic data sources, in particular linked open data, are represented by very large graph structures, typically comprising billions of individual statements (e.g. about 30 billion triples in LOD in 2011). In contrast to the conventional Web where complete documents are retrieved, the situation is different for semantic data. Users searching and exploring the data need find suitable starting points for their exploration and need to exploit the semantic relations to link them to related information.

Making sense of the data often requires filtering or aggregating them from different perspectives to extract those parts of the data graph that are relevant to the question at hand. Essentially, this means that users must be provided with tools that allow them to construct individual, potentially complex views on the data that can be used for further exploration. Due to the close integration of searching, exploring and visualizing the data, we call this process "constructive exploration". An example of this approach are the aforementioned FacetGraphs where users construct an arbitrarily complex graph of concept nodes that contain lists of instances of the respective concept. Items can then be filtered over direct or even indirect connections to other concepts, allowing the user to change ("pivot") the direction of filtering or apply several filters simultaneously.

We see this as an example of a much wider range of possibilities where users can compose and configure visualizations or user interfaces in general to suit their individual task and information needs with respect to semantic data resources. Future research will need to address suitable UI metaphors, interaction tools and implementation approaches to create such flexible and user-driven visual environments for exploring semantic data.

Dynamic Contextualization

In context-adaptive systems and user interfaces, there still exists a considerable gap between theoretical considerations and actual technical solutions. In one of the most frequently cited papers on context, Dey and Abowd (2000) argue that any information that characterizes the situation of some entity can be considered as context. However, concrete examples of adaptive systems mostly use a much more restricted view, restricting context to a priori defined factors such as location, time or the device used. Broad notions of context are hard to operationalize. To reduce these limitations, we argue that context-adaptivity should be considered as a process of *dynamic contextualization* in which the system offer means for determining relevant context on the fly when the decision about some adaptation is about to take place (Hussein et al. 2014).

Semantic models and data can play an important role in achieving such dynamic contextualization processes. Dynamic contextualization can be operationalized by a process that comprises the following steps: 1. Representing long-term user information and knowledge about the domain and potential context factors as semantic models and data. 2. Sensing potentially relevant information about the user's current state. 3. Dynamically identifying contextually relevant elements, e.g. querying the semantic database 4. Reasoning (for instance inferring recommendations) based on the extracted context. These four steps roughly correspond to a four layer model proposed in previous work of one of us (Haake et al., 2010), where a semantic model represents the knowledge base used for subsequent contextualization steps. In this framework, semantic models provide a solid and comprehensive basis for adapting UIs, using the reasoning mechanism most suitable for the current adaptation problem.

While we have explored this approach successfully in relation to context-aware hybrid recommenders and consider it promising, there are still many open research issues. Open questions exist, in particular, with respect to processing and abstracting sensed context, as well as to integrating external, sensor-based context with user models that capture behaviour or preferences. Finally, suitable reasoning mechanisms as well as meaningful and usable adaptations are required .

Transparent Integration of Semantic Data

While interacting with semantic data directly offers many opportunities for the targeted retrieval or exploration of facts and relations, in many cases it is desirable to integrate semantic data with conventional Web content to create more dedicated, application-specific user interfaces. Although this has been the subject of some research, there are still obstacles that prevent developers from making use of this option. There are issues related to architectural aspects as well as to the tools and languages used for doing so. Furthermore, there are currently no tools that would be usable by end users for integrating semantic data, e. g., in low-threshold publishing tools such as blogs or wikis.

To overcome some of these obstacles, we propose the concept of Semantic Widgets that shield users from the complexity of the RDF query language (SPARQL3) and facilitate the construction of Web applications that use and integrate semantic data. To realize the concept, we have been working on SemwidgJS (Stegemann & Ziegler, submitted), a JavaScript library for displaying Linked Open Data through Semantic Widgets. SemwidgJS can be integrated in almost any standard HTML webpage and handles the querying, processing and displaying of semantic data. While existing libraries sharing a similar goal only comprise widgets for information visualization purposes, SemwidgJS also features widgets for typical UI elements such as labels, links, and text input fields. To make querying semantic data easy, SemwidgJS supports its own simplified, path-based query language – SemwidgQL which is used as an alternative to standard SPARQL. We envisage that by providing suitable tools, users can be enabled to define the queries handled by the widgets in a completely visual way, combining exploration of the data with the definition of (reusable) widgets that can be embedded in standard Web pages.

Beyond this specific example, we see a range of research issues that need to be solved to bring semantic techniques closer into the hands of end users. In terms of architecture, server-side and client-side techniques need to be further investigated. SemwidgJS widgets, for example, are completely processed in the client, Web authors therefore only need to include suitable widget mark-up in their HTML code without the need to have specific functionality available on the server side. Integrating semantic Web services in an easy, user-definable manner is also an area where more research will be needed to move beyond presentation-oriented applications and to make mash-ups of Web content and semantic data fully interactive. Eventually, this requirement may change the overall architecture of interactive applications, possibly resulting in new solutions beyond current practices of service-oriented systems.

CONCLUSIONS

Semantic Web techniques offer a high potential for new UI styles and capabilities, and for changing the way interactive systems are designed and built. Yet, many research challenges still lie along this way. Paradigms and metaphors for interacting with semantic data, for instance must be further investigated to increase usability. More tightly integrating semantic user and context models could lead to more effective and transparent adaptations. Finally, we believe that semantic techniques could empower end users to define and compose (without programming) their own personalized applications and user interfaces.

REFERENCES

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. Scientific American, (May 2001).

Berners-Lee, T., Hollenbach, J., Lu, K., Presbrey, J., Prud'ommeaux, E. and Schraefel, m.c. Tabulator Redux: Browsing and writing Linked Data. In Proc. WWW 2008 Workshop: LDOW, (2008).

Bizer, C., Heath, T., Kingsley, I., & Berners-Lee, T. (2008). Linked data on the Web. In Proceedings WWW 2008, pp. 1265-1266.

Brophy, M. OWL-PL (2010). A Presentation Language for Displaying Semantic Data on the Web. Master's thesis, Lehigh University.

Dey, A. K., & Abowd, G. D. (2000). Towards a Better Understanding of Context and Context-Awareness. In Proceedings of the CHI 2000 Workshop on the What, Who, Where, When, and How of Context-Awareness. The Hague, Netherlands: ACM Press.

Haake, J., Hussein, T., Joop, B., Veiel, D., & Ziegler, J. (2010). Context modeling for adaptive collaboration. Int. Journal of Cooperative Information Systems, 19, 71-120.

Heim, P., Ertl, T., & Ziegler, J. (2010). Facet Graphs: Complex semantic querying made easy. In Proceedings ESWC 2010. Berlin: Springer.

Heim, P., Hellmann, S., Lehmann, J., Lohmann, S., & Stegemann, T. (2009). RelFinder: Revealing Relationships in RDF Knowledge Bases. In Proceedings. 3rd Int. Conf. on Semantic and Media Technologies (SAMT). Berlin: Springer.

Hussein, T., Linder, T., Gaulke, W., & Ziegler, J. (2009). Context-aware recommendations on rails. In CARS '09- Workshop on Context-Aware Recommender Systems (in conj. with 3rd ACM Conf. Recommender Systems).

Hussein, T., Linder, T., Gaulke, W., & Ziegler, J. (2014). Hybreed: A software framework for developing context-aware hybrid recommender systems. User Modeling and User-Adapted Interaction, 24(1-2), 121–174.

Huynh, D. and Karger, D. (2009). Parallax and companion: Set-based browsing for the Data Web. Proceedings WWW 2009.

Kobilarov, G. and Dickinson I. Humboldt: Exploring Linked Data. In: Proc. WWW 2008 Workshop: LDOW, (2008).

Pietriga, E., Bizer, C., Karger, D. R., & Lee, R. (2006). Fresnel: A Browser-Independent Presentation Vocabulary for RDF. In ISWC '06: Proceedings of the 5th International Semantic Web Conference (Vol. 4273, S. 158–171). Springer.

Stegemann, T., Ziegler, J., Hussein, T., & Gaulke, W. (2012). Interactive construction of semantic widgets for visualizing Semantic Web data. Proceedings 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS 2012), Copenhagen.

Stegemann, T., & Ziegler, J. (2014, to appear). SemwidgJS: A semantic widget library for the rapid development of user interfaces for linked open data. In AST 2014 - 8th International Workshop on Applications of Semantic Technologies (in conjunction with Informatik 2014, Sept. 22-26, Stuttgart). Berlin, Heidelberg: Springer.