GUEST EDITORIAL PREFACE

Special Issue on Visualisation of and Interaction with Semantic Web Data - Part I of II

Matthew Rowe, School of Computing and Communications, Lancaster University, Lancaster, England, UK

Aba-Sah Dadzie, Department of Computer Science, The University of Sheffield, Sheffield, South Yorkshire, England, UK

PROBLEM SETTING AND CONTEXT

The semantic web's uptake and recognition has grown in recent years. Initiatives such Google Rich snippets, Facebook Open Graph and Google Knowledge Graph, are all examples of the enrichment of data in large-scale systems with semantics to the benefit of providing more advanced services to end users. Coupled with such initiatives has been the increasing size of the linked data cloud, standing at an impressive 31 billion triples in September 2011. Such large volumes of structured data require approaches that harness the breadth of its information content and enable effective, intuitive reuse by technical users outside the semantic web community and by so-called lay users. Addressing the issue of consumption by the latter would enable more mainstream adoption of what is without doubt very valuable semantic web data in a variety of applications and scenarios (e.g. discovering local government spending, fact finding in exploratory search, identifying new bands to listen based on current interests). However making sense of semantic web data is hindered by users' lack of knowledge concerning semantic web technologies; where questions such as 'what is an ontology?' and 'what is RDF?' are commonplace. There is therefore a clear need for tools and approaches that can harness the wealth of information available, but without requiring expert knowledge of semantic web tools and technologies.

The rise in semantic web data production

means that new datasets are frequently being published about a variety of domains. Making sense of and understanding such datasets is possible through the provision of dataset documentation that explain what the dataset contains and what can be found within it, however the provision of real insights are only truly afforded through exploration of the data, especially as it grows in size and complexity. This, in turn, often requires querying the dataset to ascertain information about what is stored in it – a stumbling block for non-technical and non-domain experts, especially over RDF. Research however shows that appropriately designed data overviews provide more intuitive and accessible options for exploration, especially when these harness interactive visualization. Such representations are able to harness powerful human perception, resulting in an increase in comprehension of data structure and content, leading in turn to more effective data reuse and linking between datasets.

The semantic web is built upon the notion of shared understanding enabled through formal semantics. Such semantics are defined through the use of ontological models that capture the concepts that exist in a given domain and how they are related to one another. In a similar vein to semantic web datasets, which are (typically) built on an ontological model, ontology reuse is enabled when one understands the intricacies of the ontology defined through documentation of the ontology. However building this documentation can be time-consuming and tedious should an ontology go through several revisions prior to a final version. Additionally, the collaborative nature of ontology development is such that several engineers may be altering and evolving an ontology at once; in such situations understanding the collaboration process and how agreement is reached is important for future ontology engineering efforts. In both cases, ontology documentation and collaborative design, methods are needed to visualise the information describing either the dynamics of an ontology and how it is structured, or the collaboration process through which it was built.

This issue of the International Journal of the Semantic Web and Information Systems is the first in a two-part special issue on the visualisation of and interaction with semantic web data. We hope that the contributed work will buoy visualisation and interaction efforts in the semantic web community, towards making semantic web data more easily understandable and usable to non-semantic web users, thus enabling mainstream adoption of the community's efforts in publishing such a wide range of structured data. In this first part of the special issue we have included three papers, which we summarise below. The second part of the special issue will contain three more papers on the topics of visualisation of and interaction with semantic web data.

Semantic web datasets that contain many triples linking disparately typed resources present a challenge to users, lay and technical, trying to obtain an understanding of data content and structure. Existing approaches to this problem involve formulating queries (typically in SPARQL) to ascertain information about entities and relations in a dataset, and what properties and values are contained in the data. Formulating effective queries to extract this information can be a time-consuming and laborious process for technical users; for non-technical and non-semantic web experts this is at best very difficult, and very often, impossible. The first paper in this special issue, entitled 'From Overview to Facets and Pivoting for Interactive Exploration of Semantic Web Data' by Brunetti et al. tackles this challenge using facet-based overviews of semantic web, data, complemented by pivots, to aid navigation through and comprehension of data content. The authors present the tool Rhizomer, a semantic web data publishing and browsing tool, whose design builds on research from the domain of Information Architecture. Rhizomer provides support for obtaining an overview of datasets, and for navigation using facets and filters. An evaluation of Rhizomer with lay users demonstrates the effectiveness of the approach in enabling comprehension of semantic web data by non-technical users. The

conclusions highlight a key finding that could impact querying and data interaction methods in the semantic web domain: facets, though well known and proven to be useful for navigating through multi-faceted data such as semantic web data, are limited when relations are asymmetric. In such cases pivots may provide a more intuitive method for bridging indirect relationships, improving navigation through complex data.

Ontologies are essential elements of the semantic web stack, by offering a common understanding of some domain's interpretation. Effective reuse of ontologies relies on the documentation of classes and relations for human interpretation. However, the evolution of ontologies through versions requires the re-writing of documentation; as a consequence ontology engineers often wait until completion of an ontology to provide documentation. Automatic documentation generation therefore plays a vital role in the timely provision of ontology documentation. The second paper in this special issue, by Peroni et al., 'Tools for the automatic generation of ontology documentation: a task-based evaluation', contrasts the effectiveness and usability of three tools for generating ontology documentation: LODE, Parrot, and OWLDoc-based Ontology Browser. The authors employ two task-based evaluations to assess performance and usability, followed by a grounded analysis: (1) navigating through an ontology to complete a given task – to assess tool efficacy, efficiency (time to complete the tasks), user satisfaction; and (2) a test of the effect of ontology size on tool performance (using three ontologies of differing sizes for the tasks). Each tool displayed particular advantages, and the authors found expertise influenced the ways in which participants used the tools to perform the evaluation tasks. On average, however, the authors found the LODE tool to be considered the most usable and useful, after analysing the results across the two tasks based on tool efficacy, efficiency and user satisfaction.

The collaborative process of ontology engineering is enabled through tools such as Collaborative Protégé, which provide logs of collaboration activities between ontology engineers and other stakeholders, describing who performed what task and when. Analysing these logs allows one to assess and understand the collaboration process, and track quality-related issues such as the synthesis of contributions. However, while tools exist for visualising ontologies, there is much less support for visualising the process of engineering ontologies. With the increase in size and complexity of an ontology comes a large number of decisions made about and actions on its components and hence, the number of logged actions. Visualisation methods are therefore very useful in trying to make sense of the logs, in order to correctly interpret individual and collaborative actions. The third paper in this special issue, 'PragmatiX: An Interactive Tool for Visualizing the Creation Process Behind Collaboratively Engineered Ontologies', by Walk et al., describes the tool Pragmatix, that visualises pragmatic aspects of collaborative ontology engineering. The authors provide three network-visualisation views for exploratory analysis of the logs: concept network visualisation (to monitor generalisation and specialisation of concepts), author network visualisation (to identify individual author actions and commonly authored concepts and properties), and property network visualisation (to monitor modification of properties and the relationships between different properties). These are complemented by statistical analysis, using chart and tables views. Walk et al. report a summative evaluation of Pragmatix with domain experts in bio-medicine, that provides information on the effectiveness of the prototype for supporting the collaborative ontology engineering process.

We would like to thank Amit Sheth for inviting us to edit a special issue of the journal and for his help throughout the process. We would also like to thank the following reviewers who worked hard to ensure that we received high quality reviews for all submitted work.

Matthew Rowe
Aba-Sah Dadzie
Guest Editor
IJSWIS