

SDSW' 2014 Panel Report

Research challenges and Avenues in Surfacing the Deep and the Social Web

After the paper presentations there was a panel where the audience, authors, and workshop organizers discussed research challenges and avenues. To that effect, four questions were used to guide the conversation. The consolidated answers are presented below.

Q1 - By many the Semantic Web (SW) has not managed to reach its original goal. Nowadays that Big Data, Cloud computing and Social Networks have been added into the picture, do you think that this goal became closer or further?

The SW roadmap has been established in 1998. With respect to its original goal, some did argue that the SW managed to be closer, in promoting RDF as a kind of unified model; although there was some shift according to two directions: from documents to metadata, and from data to resources.

With respect to the Big Data impact, shifting between communities makes it difficult to handle big data. There is more and more knowledge embedded in the data, and other techniques such as machine learning should be used to assist the data analysis process. There is also a difference between engineering and science, industry and academia: although (most) enterprise tools are SW-limited, there may be some exceptions such as the Google Graph Knowledge. Finally there is also a multilingualism issue: as an example, counting for the same entity in DBpedia may be different from German to French.

Q2 - Do you see the privacy issue playing a role in your approach, or do you consider it completely orthogonal?

It was not clear that the privacy did play a significant role according to the papers presented at the workshop.

However, the discussion highlighted some issues and candidate solutions:

- *issues: privacy leaks (notably with open data), need for a semantics of privacy, security vs. privacy.*
- *solutions: data anonymization, graph structure encryption, ontology watermarking.*

Q3 - The ability to effectively express the information that a user is looking for is of paramount importance. Recently, people get further and further away from SQL and other traditional languages. Keyword query seems to be a favorite solution, yet with a lot of limitations. What is your opinion about modern or new forms of querying techniques?

The real issue is to help users express or anticipate their needs. Search engines are building Knowledge Graphs but most of the time the user has the knowledge and does not know how to express it: exploratory search solutions are being proposed to tackle this issue. Also, there is a need for visualization-based solutions.

Q4 - The deep web looks like a collection of structured sources. How do you see the integration of these data with the social and the document web data into one unified framework?

The discussion about RDF (as a unified model) was brought up again. Some did argue that we could almost investigate everything: see, for example, the social tagging. Others did argue that we may need some complex multi-layered (graph) model to handle various dimensions such as temporality and spatiality: as an example, community detection is temporality dependent. Hence, there is a need for a multidimensional analysis on a multilayered model, and more investigation is needed with respect to a unified model/framework. Furthermore, it is clear that there is no single solution for the semantic web, rather than a collection of tools and techniques orchestrated together. So, it is not about simply the data model. Data can be represented flexibly. However, it is about building applications that can communicate between them.