Web-Information Systems 2.0? Social, Economical, and Technological Challenges

Pieter De Leenheer and Robert Meersman

Semantics Technology and Applications Research Laboratory (STARLab)
Vrije Universiteit Brussel
Pleinlaan 2, B-1050 BRUSSELS 5, Belgium

The Seminar

In less than two decades, computing and information technology in particular have undergone a fundamental change. The World Wide Web successfully evolved from a static web of "one-way published" media resources towards a knowledge sharing platform that allows on-line communities to emerge all around, massively share ideas, and create knowledge in a usually self-organising manner. This "second-generation Web" (Web 2.0) is manifested by an explosion of rich internet applications and technologies being developed and shared at little or no cost, including social applications like bookmarking, lightweight folksonomies, blogs, and wikis.

The emerging range of Semantic Web and other open technologies, pulled by the networked economy, and pushed by the exponential growth of computing power and bandwidth, promises an increase in scale and maturity of this Social Web phenomenon, achieved through collaboration and integration within and between different and diverse communities. Some visionairs hypothesise about a new wave of Internet revolution, which would ultimately take us to the so-called Internet of Services or Web 3.0, where communities will consist of a mix of human and software agents "crowdsourcing" services from each other, orchestrated by the goals and needs of the community. Through advanced mobile technology, these services would become ubiquitously available on domestic devices other than personal computers, making the latter disappear and become obsolete for end users.

The Web has become pervasive and interacts deeply with numerous business processes that until now have been the internal concern of enterprises. An attempt to understand the rules for success on that new platform has leveraged the long tail that revolutionises internet business, and empowers a creative minority "at the edge" of the crowd to push forward innovative products that have a high potential to become mainstream. Such a novel view on business ecosystems could benefit particularly the success of SMEs that is dependent on finding the right partners, manage these relationships, and ultimately extend the reach of their particular products and services.

When extrapolating the current pace of technological development, it seems that the transformation of the Web from a network of separately siloed applications and content repositories to a more seamless and interoperable ecosystem is at hand. However, we are facing a universal limit on the total amount of information that can be stored and processed in the future, which is expected to put an ultimate time constraint on Moore's

Law. Furthermore, although XML (and its successor RDF(S) and OWL) is becoming a de facto format for structuring and exchanging data and services, the degree of explicit shared semantics in it is marginal. Therefore, massively distributed information services and semantic reconciliation of heterogenous data silos are key drivers that will determine the Web's ability to carry out the envisioned services on the underlying knowledge without the need for human intervention.

Ontologies, being formal, computer-interpretable specifications of shared conceptualisations of the worlds under discussion, are an important lever by providing shared resources of semantics. At first sight ontologies may look similar to (extended) data models, enhanced with business rules and constraints, but the latter are in general intended for a particular, a priori known, application that their author had in mind for his/her enterprise. A shared, non-application specific resource must indeed be the result of community-based (rather than single organisation-based) agreement on the concepts, relationships, events, rules and processes occurring in that domain. Agreements imply (often virtual) communities of users, experts and/or designers that collaborate towards that shared, evolving, and formal understanding.

Architecting internet systems à la Web 2.0 will require a paradigm shift that goes beyond technological excellence. Paving the way will start with radical, multi-disciplinary reconsideration of the ecology of its first-class citizens, namely (meta-)data, processes, services, and predominantly its owning communities. Current approaches tend to reduce communities to a large set of contributing individuals only, while ignoring their legitimate participation in the community. Just as database schemas and ontologies respectively took the data and much of its semantics out of applications, the proposed approach goes further by externalizing also the semantics of the (user) communities themselves. Social Web and Web 2.0 are providing fresh insights into these issues, and vice versa. For example, Web 2.0 made us realise that unleashing ICT on massive communities results in a rapid emergence of invaluable common sense in a self-organising manner. Such a manifestation of wisdom of crowds was never thought to be conceivable with traditional AI.

On the first day, this workshop features a number of invited talks by senior experts in the field, each addressing relevant challenges in next -generation web-information systems. The second day of the workshop is dedicated to extended discussion. The discussion will be initiated by a panel consisting of the five invited speakers. In the afternoon, there will be an opportunity to brainstorm about the identified challenges and opportunities.

Location

Vrije Universiteit Brussel, Campus Etterbeek, Pleinlaan 2, B-1050 BRUSSELS
The workshop takes place in Building E Room 0.06. Travel Directions:
http://www.vub.ac.be/english/infoabout/campuses/index.html

The Programme

Day 1: 17 September

- 9:00: Welcome & Opening by Prof. Dr. Robert Meersman
- 9:30: **Information Retrieval and Data Mining in Social Bookmarking Systems**
  Invited Speaker: Prof. Dr. Gerd Stumme
  Abstract: Social bookmarking systems allow users to organize and share bookmarks, photos, etc. on the web. The reason for the success of these systems lies mainly in the fact that no specific skills are needed for publishing and editing. As these systems grow larger, however, the users feel the need for more structure for better organizing their resources. This poses new challenges for the research areas of information retrieval and knowledge discovery. The underlying data structure of social bookmarking systems are so-called folksonomies, which can be considered as tri-partite hyper-graphs. One approach is thus to transfer existing graph algorithms to this new kind of structure. In the talk, I will first present our social bookmarking system BibSonomy, which supports the scientific community in organizing and sharing bookmarks and lists of publication references; and which serves as a platform for our experiments with Information Retrieval and Knowledge Discovery algorithms. In the second part of the talk, I will present selected research results.
  For more information:
  - http://www.kde.cs.uni-kassel.de/stumme
  - http://www.bibsonomy.org/
- 10:30: Coffee Break
- 11:00: **Ontological Foundations for Community Evolution**
  Speaker: Mr. Pieter De Leenheer
  Abstract: Research in ontology management has reached a certain level of maturity, however, there is still little understanding of, and technological support for, the methodological and evolutionary aspects of ontologies as resources. Yet these are crucial in distributed and collaborative worlds such as the Semantic Web, where ontologies and their communities of use naturally and mutually co-evolve. Through a deep understanding of the real-time, community-driven, evolution of so-called ontologies, knowledge-intensive system can be made operationally relevant and sustainable over long periods of time. Such a paradigm shift in knowledge-intensive and community-driven systems would affect knowledge sharing and communication across diverse communities in business, industry, and society. In this paper, we give an overview of the practical and theoretical challenges and limitations, and based on that introduce an ontological and methodological foundation for community evolution processes. For more information:
  - http://www.pieterdeleenheer.be
- 12:00: Lunch
- 13:30: **Ontology Economics**
  Invited Speaker: Prof. Dr. Martin Hepp

Abstract: The state of the art in ontology building and modeling in general widely disregards the fact that the behavior of rational actors in the modeling process will be driven by individual economic incentives. Quite obviously, creating an ontology requires resources (e.g. labor) from the contributing individuals. Now, it is for many domains a completely open research question whether the gain in automation, to be made possible by the availability of a respective ontology, outweighs the resources necessary for creating the ontology. Even worse, even if the overall benefit of an ontology over its lifetime exceeds the costs for creating the ontology, it may still not be created if it is not economically feasible for each individual required to contribute. In this talk, I will summarize the research questions of this novel field of Ontology Economics and present preliminary results. For more information:

- http://www.heppnetz.de
- http://www.unibw.de/ebusiness/

– 14:30: **Data Quality, Compliance, Data and Web mining for Business Intelligence**

Invited Speaker: Elizabeth Chang

Abstract: In this seminar, professor Chang through 2 case studies, introduce the importance of Data, data Quality, Compliance and Data and Web mining for Business Intelligence. Data is one of most valuable asset for any organization. Data Quality is vital to BI implementation success. The national data audits have determined that 15% of the data in the typical customer database is inaccurate, and poor quality information have been measured at the cost of 20% of the organizations revenue per year. Data quality is about data that is accurate, reliable and trusted. Improve data quality, improves the quality of the decisions and improves the effectiveness of the decisions as to its customers and stock holders. Compliance defines the scope for Business Intelligence. It revitalizes operational efficiency and competitive advantage. It is about trust, credibility and relationships. It contributes to cost savings and operational efficiency. It focus on the strategies, policies, processes, procedures, and enforcement and accountable for timely analysis, accurate and updated information, protected privacy and confidentiality, controlled access and authorized data views, security and attested proof of above. The Web is a service platform in the new digital economy, has radically changed business models and transformed communities. Organization today shall offer its services through internet, mobile devices and networks and use this as a means to automate its operation. It is all about allowing people (customers) and partners to access services ubiquitously at anywhere and 24-7. In addition, acquisition of market information, knowing what customers are thinking, what they want, how to serve to them, what drives them. This is part of the business intelligence and it is priceless! Professor Chang will also talk about a range of Data quality improvement techniques, data warehouses strategies, risk and compliance templates and frameworks, data mining solutions, BI implementation maps and measures. For more information:

- http://www.fit.cbs.curtin.edu.au/ chang

– 15:30: Coffee Break

- 16:00: **Ontology Evolution using Social Networks and Web 2.0**
  Invited Speaker: Prof. Dr. Tharam Dillon
  Abstract: to be provided
- 17:00: Concluding remarks
- 17:30: Reception

Day 2: 18 September

- 9:00: Recap Day 1 and Opening Day 2 by Mr. Pieter De Leenheer
- 9:30: **Panel Discussion** with panel members: Prof. Dr. Tharam Dillon, Prof. Dr. Gerd Stumme, Prof. Dr. Martin Hepp, Prof. Dr. Robert Meersman, Prof. Dr. Elizabeth Chang
- 12:00: Lunch
- 13:30: Continuation of the Discussion and Brainstorm
- 15:30: Coffee Break
- 16:00: General Conclusions
- 16:30: End of the Workshop