Undefined 0 (2010) 1–0
IOS Press



# Towards the Ubiquitous Web

**Editor(s):** Krzysztof Janowicz, Pennsylvania State University, USA and Pascal Hitzler, Wright State University, USA
**Solicited review(s):** Marta Sabou, MODUL University Vienna, Austria and Boyan Brodaric, Geological Survey of Canada, Canada
**Open review(s):** none

Andreas Hotho [a,b] and Gerd Stumme [a,c]

[a] *Knowledge & Data Engineering Group, Department of Mathematics and Computer Science, University of Kassel, Wilhelmshöher Allee 73, 34121 Kassel, Germany*
*http://www.kde.cs.uni-kassel.de/*

[b] *Data Mining and Information Retrieval Group, Department of Mathematics and Computer Science, University of Wuerzburg, Am Hubland 97074 Wuerzburg, Germany*
*http://www.is.informatik.uni-wuerzburg.de/*

[c] *Research Center L3S, Appelstr. 9a, 30167 Hannover, Germany*
*http://www.l3s.de/*

**Abstract.** Today, we observe the amalgamation of the Social Web and the Mobile Web, which will ultimately lead to a Ubiquitous Web. The integration and aggregation of the different kinds of available data, and the extraction of useful knowledge and its representation has become an important challenge for researchers from the Semantic Web, Web 2.0, social network analysis and machine learning communities. We discuss the Ubiquitous Web vision, by addressing the challenge of bridging the gap between Web 2.0 and the Semantic Web, before widening the scope to mobile applications.

Keywords: Ubiquitous Web, Semantic Web, Web 2.0, Data Mining, Machine Learning, Social Network Analysis

## 1. Introduction

Highly popular user-centered applications such as blogs, social tagging systems, and wikis have come to be known as "Web 2.0" or the "Social Web" [6]. At the same time, mobile phones became more and more powerful and are equipped with more and more sensors, giving rise to Mobile Web applications. Today, we observe the amalgamation of these two trends, leading to a Ubiquitous Web, whose applications will support us in many aspects of the daily life at any time and any place. The integration of the different kinds of available data, their integration and aggregation, and finally the extraction of useful knowledge and its representation has become an important challenge for different research communities, since it requires the confluence of previously separated lines of research. Consequently, the last years have seen increasing collaboration of researchers from

the Semantic Web, Web 2.0, social network analysis and machine learning communities. Applications that use these research results are achieving economic success. Data now become available that allow researchers to analyze the use, acceptance and evolution of their ideas.

In this position paper, we will discuss the Ubiquitous Web vision in two steps. First, we will address the challenge of bridging the gap between Web 2.0 and the Semantic Web, before widening the scope to mobile applications.

## 2. Bridging the Gap Between Web 2.0 and the Semantic Web

A major reason for the immediate success of Web 2.0 systems is their high ease of use. The result is that the "wisdom of the crowd" and the wisdom of the experts are converging. The online encyclopedia Wikipedia, for instance, reaches (and in some areas even surpasses) the quality of traditional dictionaries [3]. We anticipate that wikis, resource sharing systems, and blogs are only the first appearances of an emerging family of web cooperation tools. These sites do not only provide content but also generate an abundance of weakly structured metadata. A good example is tagging. Here, users add keywords from an uncontrolled vocabulary, called tags, to a resource. Such metadata are easy to produce, but lack any kind of formal grounding, as used in the Semantic Web.

The Semantic Web complements the bottom-up effort of the Web 2.0 community in a top-down manner. Its central point is a stronger knowledge representation, based on some kind of ontology with a fixed vocabulary and typed relations [5]. Such a structure is typically something users implicitly have in mind when they provide their content in Web 2.0 systems. However, for further use, this structure is hidden in the content and needs to be extracted. In the Semantic Web community, such approaches are known as Ontology Learning [1]. Techniques to analyze network structures or weak knowledge representations, such as those found in the Web 2.0, have also a long tradition in different other disciplines, like social network analysis, machine learning and data mining. These kinds of automatic mechanisms are necessary to extract the hidden information and to reveal the structure in a way that the end user can benefit from it. Using established methods to represent knowledge gained from unstructured content will also be beneficial for the Web 2.0 in that it provides Web 2.0 users with enhanced Semantic Web features to structure their content.

Besides the application of Semantic Web technology, it may also be beneficial to consider more light weight knowledge representations, since not always approaches with strong formal semantics are needed. One example are statistical representations, such as association rules, tag similarity measures, and similarity measures in search engines or recommender systems. A careful analysis of the intended application will decide the way to be followed.

The main research question can be summarized as follows: *How will current and emerging Web 2.0 systems support untrained users in sharing knowledge on the Web within the next years?* The scientific challenge is to develop "minimal-invasive" and scalable techniques for large, web-wide spread user communities for knowledge sharing. While knowledge acquisition and management has a long research history, the new aspect of the Web 2.0 is a) the real large number of users who are willing to share their knowledge but b) who are very selective in participating and will stop their cooperation soon if the barriers are set too high. An import requirement is thus how to build, from the uncoordinated input of many people, where each individual is providing very little and/or unstructured input only, a shared knowledge space which allows for similar benefits as those usually promised for approaches with one central, well-designed, heavyweight ontology. This representation will probably not be presented to the users, as the interaction has to be kept as simple as possible, but will be the basis for the systems' enhanced functionalities.

## 3. The Ubiquitous Web

The emergence of ubiquitous computing [7] has started to create new environments consisting of small, heterogeneous, and distributed devices that foster the social interaction of users in several dimensions. Similarly, the upcoming Social Semantic Web also integrates the user interactions in social networking environments. For instance, nowadays modern smartphones allow everyone to have access to the WWW at every place and at every time.

At the same time, these systems are equipped with more and more sensors. Typical sensors in today's smartphones are measuring proximity, ambient light, acceleration, loudness, moisture, geographic north. Furthermore, access to the most prominent Web 2.0 platforms – in particular Facebook, Flickr, Youtube – is frequently pre-installed by the vendor. This example shows that the worlds of WWW, Web 2.0, the Mobile Web, and sensor technology are rapidly amalgamating. Going even one step further, we assume the rapid convergence of the Ubiquitous Web with the Internet of Things (cf. [2]) — more and more, the real world that is surrounding us will have its digital counterpart.

Applications in the Ubiquitous Web will thus rely on a mix of data from sensors, social networks and mobile devices. These data need to be integrated, aggregated, and analyzed by means of Data, Text, and Web Mining techniques to all for semantic and/or statistical representations of knowledge, which will then fuel the ubiquitous applications.

Mining in ubiquitous and social environments is thus an emerging area of research focusing on advanced systems for data mining in such distributed and network-organized environments. It also integrates some related technologies such as activity recognition, Web 2.0 mining, privacy issues and privacy-preserving mining, predicting user behavior, etc. (cf. [4])

In typical ubiquitous settings, the mining system can be implemented inside the small devices and sometimes on central servers, for real-time applications, similar to common mining approaches. However, the characteristics of ubiquitous and social mining are in general quite different from current mainstream data mining and machine learning. Unlike in traditional data mining scenarios, data does not emerge from a small number of (heterogeneous) data sources, but potentially from hundreds to millions of different sources. As there is only minimal coordination, these sources can overlap or diverge in any possible way.

Semantic Web technology can bridge the gap between all kinds of information independent of its source and its origin and can be used as a starting point to put everything together. The real world information gathered by sensors will be used by applications running on mobile devices and will be connected with the information of their users from the social web. Semantic Web technology may become the right knowledge representation for connecting these worlds.

## 4. Conclusion

Today, we see the first steps towards an integration of the Social, Mobile and Semantic Webs. This path allows for exciting challenges for researchers of different communities. New insights provided by machine learning and social network analysis techniques will lead to a new type of knowledge. We envision that research in this area will be of growing interest, as the automatic extraction of knowledge from weakly structured sources contributed by a huge mass of users and the combination with structured knowledge will be an important basis for the Semantic Web. It will lead to a broad range of new applications, which allow for combining knowledge of different types, levels and from different sources to reach their goals. The upcoming Ubiquitous Web is one target application area which will benefit from the newly integrated knowledge.

## References


[1] P. Cimiano. *Ontology Learning and Population from Text: Algorithms, Evaluation and Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[2] S. Dodson. Internet of things. *The Guardian*, 2003.

[3] J. Giles. Internet encyclopaedias go head to head. *Nature*, 438(7070):900–901, 2005.

[4] M. May, B. Berendt, A. Cornuéjols, J. Gama, F. Giannotti, A. Hotho, D. Malerba, E. Menesalvas, K. Morik, R. Pedersen, L. Saitta, Y. Saygin, A. Schuster, and K. Vanhoof. Research challenges in ubiquitous knowledge discovery. In *Next Generation of Data Mining (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series)*. Chapman & Hall/CRC, 1 edition, 2008.

[5] S. Staab and R. Studer, editors. *Handbook on Ontologies*. Springer, Berlin, 2. edition, 2009.

[6] O. Tim. What is web 2.0? design patterns and business models for the next generation of software. http://oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html, 2005.

[7] M. Weiser. Ubiquitous computing. *Computer*, 26(10):71–72, 1993.