

Research on collaborative information sharing systems

Davide Eynard

Politecnico di Milano
Dipartimento di Elettronica e Informazione
Via Ponzio 34/5, 20133 Milano, Italy
eynard@elet.polimi.it

Collaborative systems are systems designed to help people involved in a common task achieve their goals. They are widely used today, and they're gaining a great consensus both inside corporations and on the World Wide Web. There are many kinds of collaborative systems, such as Wikis (like Wikipedia), blogs, tag-based systems (like Flickr, del.icio.us and Bibsonomy) and even collaborative maps (as in Google Maps). One of the main reasons of this success is that, as applications are becoming more and more data-driven, spontaneous user participation adds value to a system because it helps in creating a new, unique and hard to recreate source of data [1].

The main objective of this research project is to study collaborative systems and the possibility to enhance them through semantics. The aim of a contamination between these systems and Semantic Web technologies is twofold: on one side, we think that the huge quantity of information created by the participation of many users can be better managed and searched thanks to added semantics; on the other side, Semantic Web community can exploit spontaneous collaboration to increase the amount of knowledge described through formal representations, making it available to many other applications. Between the many different collaborative systems currently available we chose a couple of families which, in our opinion, presented the most interesting open problems. On one side, we approached Wikis and their semantic extensions [2-4]. On the other side we studied tag-based systems (also called *folksonomies*), with a particular attention to *social bookmarking* web sites, highlighting their advantages and their limitations[5-9].

One of the main problems which characterizes Wiki systems is that published information is unstructured, hard to search and manage. Current research on Semantic Wikis is trying to address this problem through formal descriptions of Wiki contents. Folksonomies have limits which are mostly due to their self-moderation: lack of precision, lack of recall, gaming (that is, anyone can pollute the system intentionally with wrong information), and lexical ambiguities [8], which do not allow to easily extract meaning from tags in ways other than statistics and clustering. To address Wiki limitations, we are working on a model which uses different ontology layers to describe not only the contents, but also the context (that is, the processes and the dynamics between users inside the wiki) and the system itself. This would make the system not only more interoperable with other applications, but also more easy to shape, so it would better suit the

needs of particular *communities of practice* [10]. For what concerns folksonomies, we decided to extend them with semantics in two different ways. On one side, using ontologies to describe them in a formal way: through these *folksologies* [11, 12], it is possible to model tag-based systems allowing for interoperability on different levels (inside the single user space, within a system, or between different systems and users). On the other side, using ontologies to describe folksonomy contents rather than structure, mapping user tags inside it: this would allow users to both have a quick, bottom-up, easy to use tag space and a more formalized, top-down hierarchical view of their tags.

At the present time we have implemented a tool which maps tags from del.icio.us inside Wordnet ontology and provides a new way to browse them: this allowed us to address some of the main problems which are typical of folksonomies, such as lack of recall and lexical ambiguities. We have also developed a fuzzy model to describe tag-based systems, which allowed us to get more accurate results through advanced fuzzy queries and to formally describe properties of some particular classes of tags. Currently, we are working on a Semantic Wiki prototype which implements the model previously described, also allowing users to tag its contents and map their tags inside a domain ontology.

List of References

1. Tim O'Reilly. What is web 2.0. design patterns and business models for the next generation of software. September 2005.
2. Bo Leuf and Ward Cunningham. *The Wiki Way: Collaboration and Sharing on the Internet*. Addison-Wesley, 2001.
3. Max Völkel and Sebastian Schaffert, editors. *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics*, Workshop on Semantic Wikis. ESWC2006, June 2006.
4. Michel Buffa, Gal Crova, Fabien Gandon, Claire Lecompte, and Jeremy Passeron. Sweetwiki : Semantic web enabled technologies in wiki. In Völkel and Schaffert [3].
5. Clay Shirky. Ontology is overrated: Categories, links, and tags, 2005. http://www.shirky.com/writings/ontology_ouerrated.html.
6. Ellyssa Kroski. The hive mind: Folksonomies and user-based tagging, Dec 2005. <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/>.
7. Tony Hammond, Timo Hannay, Ben Lund, and Joanna Scott. Social bookmarking tools (i): A general review. *D-Lib Magazine*, 11, Apr 2005.
8. Scott Golder and Bernardo A. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, April 2006.
9. Hana Shepard, Harry Halpin, and Valentin Robu. The dynamics and semantics of collaborative tagging. In *Proc. of the 1st Semantic Authoring and Annotation Workshop (SAAW2006)*, 2006.
10. Etienne Wenger. *Communities of Practice. Learning, meaning, and identity*. Cambridge University Press, New York, Port Chester, Melbourne, Sydney, 1998.
11. Tom Gruber. Tagontology - a way to agree on the semantics of tagging data, 2005. <http://tomgruber.org/writing/tagontology-tagcamp-talk.pdf>.
12. Stefano Mazzocchi. Folksologies: de-idealizing ontologies, April 2005. <http://www.betaversion.org/~stefano/linotype/news/85/>.