Learning With Social Semantic Technologies - Exploiting Latest Tools

Gisela Granitzer[1], Klaus Tochtermann[123], Patrick Hoefler[1]

Know-Center Graz[1], Graz University of Technology[2], Joanneum Research[3]

Key words: *Education, Knowledge Structuring, Semantic Technologies, Semantic Web, Social Software, Social Semantic Technologies, Semantic Wiki, Semantic Blogging*

1 Introduction

Since the emergence of Web 2.0 and its easy-to-use web based applications, summarized under the term Social Software, ordinary Internet users are empowered to generate and publish content themselves (O'Reilly 04). During the last two years, it has become quite popular to externalize knowledge on the Web by using social media tools including, most of the time Wikis and Weblogs. While in the past users only consumed information, now they are actively producing content. The enormous growth rate of the blogosphere (Duarte et al. 07), the collection of all Weblogs, is a clear sign of an age of user generated content.
Even though it was only about three years ago that Web 2.0 became a trend, starting with applications such as YouTube or Flickr, it has become a common practice to utilize Social Software in enterprises (e.g. Back et al. 08) as well as in learning institutions, as mentioned in section 2.1. The value is clear: Social Software is easy to use. Because of being web-based it is accessible from everywhere. And it perfectly supports collaboration and communication which was not that simple before Social Software when monolithic, rather cumbersome systems dominated learning environments.

Of course, beside all its positive aspects, Social Software introduces several challenges. First, there are huge amounts of information that are easily generated with Social Software. The already existing information overload (Gantz 07) gets even worse. It becomes more and more difficult to process information and find relevant knowledge. Second, the generated information often is unstructured and not well interlinked, not to mention the fact that the information is distributed across different systems. Third, and this is more of an organizational issue, generated information becomes idle and is not reused. These three challenges are not only true for the public Web but also for individual organizations such as learning institutions where e.g. weblogs are heavily used during a semester but not afterwards.

Since it is not conceivable that we will overcome the mentioned challenges without the help of information technologies, we urgently need tools that help us cope with the mentioned issues. Traditional knowledge systems such as knowledge, content, or learning management systems are not appropriate for supporting dynamic everyday working and learning, because they are all too static and standardized. For that reason, knowledge discovery, transfer, and acquisition must be organized in a way that makes it easy for the user to survey the loads of often unstructured information on demand. To our estimate the solution lies in underpinning Social Software with structure resulting in Social Semantic Software. Thereby information becomes more easily accessible and better reusable.

The remainder of this article is structured as follows: Chapter 2 starts with a brief introduction to Social Software and Semantic Web, as these two concepts build the basis of semantically enabled Social Software, which is described in the same chapter. In Chapter 3 we illustrate learning scenarios where semantically enabled Social Software is applied. Concluding remarks and a future outlook are given in Chapter 4.

2 Social Software Meets Semantic Technologies

In the following we introduce Social Software, Semantic Technologies and Social Semantic Technologies. We do this from a conceptual point of view, since this article does not claim to outline technological details. Our goal is to give the reader an idea about which applications are available and how the might be applied to learning situations.

2.1 Social Software

Social Software, e.g. Wikis, Weblogs, Social Media Sharing, Social Bookmarking, Podcasting, or Instant Messaging, supports and enables interpersonal communication, interaction and collaboration and is characterized by a high level of self-organization of the users involved. The idea behind Social Software is that users produce content and make it available to others. This induces a human web, which mainly builds on user generated content. The term architecture for participation accurately describes the idea behind this. Social Software is part of the so-called Web 2.0 – often wrongly put on the same level with it – which emerged about three years ago. The term was coined by Tim O'Reilly and colleagues when they prepared a web technology conference in October 2004. Corresponding concepts, technologies and applications attracted increasing attention since then, not only in the private but also in the organizational sector. Principally, Web 2.0 rests on three pillars: content, community, and services, which resemble the eight design principles as defined by (Tim O'Reilly 04).
In learning institutions there is a variety of application scenarios: Weblogs are used as a means for supervising students who work abroad (Pauschenwein et al. 06), Wikis are used for collecting factual knowledge within a course (Ebner et al. 06), or Podcasting is used for recording and publishing lectures (Nagler et al. 08), to mention just a few.

2.2 Semantic Technologies

While, as before stated, the Web 2.0 can be associated with a human web, which in particular builds on user generated content and networks people, the Semantic Web constitutes a machine-processable web of data that is highly structured. Because of its highly formal and coherent description, this data can be processed by machines in a meaningful way. For this purpose data must be application-independent, composeable, classified, and part of a larger information ecosystem according to (Daconta et al. 03). Technologies of the Semantic Web of course need not necessarily be applied to the Web, but can also be applied to any kind of information collection. As an example, consider the competences a student will have acquired at the end of her studies. To bring her there, she must be provided with learning materials, be it scripts or websites, which meet her current competence level, which of course develops over time. Since competences are not independent of each other, they are structured according to predefined relationships. In order to provide the student with learning materials matching her competence level, the learning material would be annotated with the competence concepts.

According to (Daconta et al. 03), three problems suggest that there is a need for Semantic Web technology. First, there is information overload. The information quantity continuously

increases, but the human information processing capacity does not. For that reason, it becomes more and more difficult to find and select relevant data, which is important in professional as well as private situations. Second, usually data is stored in monolithic systems, also called stovepipe systems. This raises difficulties when it comes to sharing data across databases. Searching and finding remains restricted to the individual systems. It is the work of people to connect and integrate data which yields dissatisfying results. Third, there is the need for content aggregation. Even though it can be done on an HTML basis, namely syntactically, it is not yet possible to aggregate content based on its meaning.

Humans of course can handle these problems, since they are able to filter, infer, map, and combine content, but only on a small scale. Machines cannot do that, even though principally they would have the capacity. And this is exactly the vision of the Semantic Web. According to this vision, machines will process information in a meaningful way, with the meaning coming from a defined structure of the data. The main technologies which will empower machines to understand the meaning of data are XML, RDF, the corresponding schemata, and Ontologies.

2.3 Semantically Enabled Social Software

Social Software and Semantic Web have initially built two separate, antithetic streams, with their advocates not seeing the chance of integration. Recently, however, one can observe an increasing convergence of Social Software and Semantic Web to a Social Semantic Web, often also referred to as Web 3.0. There are two variations of the Social Semantic Web: semantically enabled Social Software and socially enabled Semantic Web. Of course, this does not only refer to the Web but to any other information collection as well. The first variation refers to the enhancement of Web 2.0 content by machine-processable semantic data. The second variation refers to the collaborative creation of structured semantic data. Even though these two variations are conceptually different, they reflect two sides of the same medal. In the following we will introduce possible semantically enabled Social Software applications [(Pellegrini & Blumauer 07), (Pellegrini & Blumauer 08), (Schaffert 06)], since our contribution has its focus on this variation.

In the case of Semantic Wikis[1], the content of a Wiki is mapped to a predefined structure which machines can »understand«. The basis is a structure consisting of concepts connected to each other by specified relations. For a better understanding, consider the following example. If the information about university courses were structured by a »*course*« »*deals with*« a certain »*topic*« and a »*lecturer*« »*holds*« a »*course*« a student searching for courses would automatically be given topics and lecturers referring to the found courses. If he would search for a concept and a relation, e.g. »*lecturer*« and »*holds*«, he would immediately find all the courses a lecturer gives. Thus, the structure allows for a very efficient and unerring search. For an overview of common Semantic Wiki features see (Schaffert et al. 06).

Weblogs also mix with Semantic Technologies in the form of Structured and Semantic Blogging (Cayzer 06). Structured blogging means that machine processable data such as geo-coordinates, contact information, calendar data or keywords enrich the code behind Weblog entries, which makes them better searchable. An example is the WordPress plugin Yahoo!Shortcuts[2] which detects named entities such as locations, persons, organisations, or products within Weblog entries and enriches these entities semantically. Users can add materials such as photos to these entities, making the information even richer. If an Ontology structures the additional data, we talk about semantic Blogging.

[1] e.g. http://ikewiki.salzburgresearch.at/ or http://www.semantic-mediawiki.org/
[2] http://shortcuts.yahoo.com/

Communities can also be supported by utilizing Semantic Technologies. Identifying and annotating actors and relations among them and laying this information down in a format suitable for the Semantic Web helps finding people with complementary or similar competences. This can be of interest when someone has to find a co-author or an expert for a joint project proposal. Also, suggestions for possibly interesting communities can be made on this basis. Another approach is followed by Twine. This software analyzes content that a user flags important during her daily work. These contents are enriched with semantic information and are interpreted as interest profiles. These profiles are matched so that recommendations concerning content, people or topics can be generated. Another well known project is Socially Interlinked Online Communities[3]. In this project an Ontology was developed including the central concepts of online communities such as user, role, or post and describing the relations among them. By this coherent description various online communities, even based on different tools such as Weblogs, Chats or Forums are connected to each other. A query would span all these communities and tools yielding e.g. all community statements matching the query.

3 Semantically Enabled Social Software in Practice

In the following we will show three scenarios about how semantically enabled Social Software could be applied to educational situations at a university. They are only scenarios since, to our knowledge, there do not exist corresponding real life applications.

3.1 Scenario 1: Semantic Wiki

3.1.1 Problem

Consider students who have to write their bachelor theses. Usually, the students get topic and supervisor assigned, then they each work on their topic and finally they hand in a report. Final reports may be archived in a central system, but all the additional information such as used publications, lecture notes or websites, contacts to colleagues, or utilised communities are lost. Also during work this information is not transparent and thus synergies cannot be generated.

3.1.2 Solution

For optimally using the bachelor thesis related information during work and subsequently, a Semantic Wiki would be an option. First, a basic structure including objects and annotated links has to be developed. Figure 1 shows how such a structure could look like. For a better understanding we kept it rather simple, however it could be extended by classes such as *course* or even *faculty* and corresponding annotated relations.

[3] http://sioc-project.org/

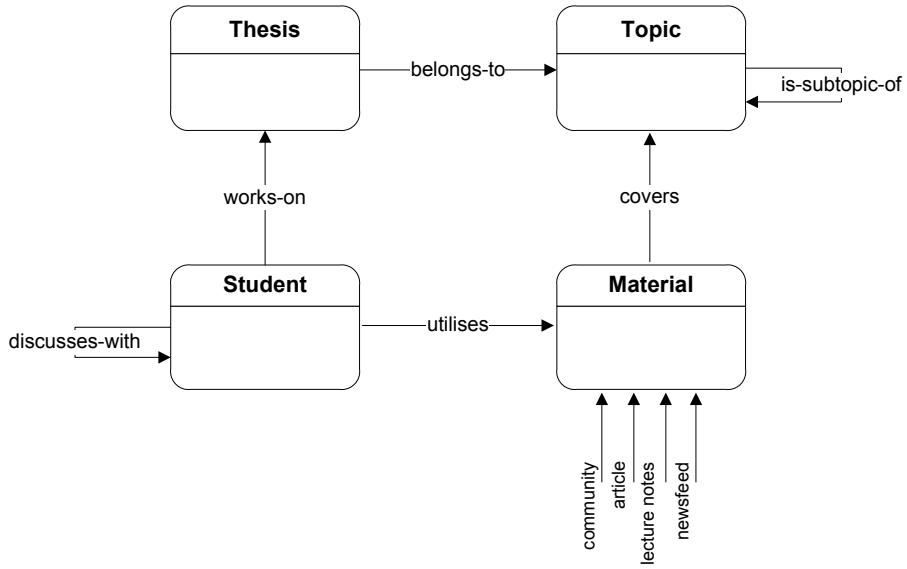

Figure 1. Structure describing knowledge objects and their associations

Students working with an accordingly designed Semantic Wiki would now enter information which is relevant in the context of their bachelor work. Of course, this information must be entered according to a given scheme, so that the relationships can be exploited to the benefit of the students. Each student would create a kind of continuously growing profile regarding her thesis. While documenting her ongoing work and results, she would link or upload materials and relate them to the topic her thesis belongs to. She would also document discussion partners. Since usually more students work on the same topic together, they can collectively write about that topic as it is commonly done in Wikis.
Of course, an interesting question is how the Semantic Wiki can be of value for the students currently working on their theses and for future students who simply want to learn about a topic. Principally, the value lies in the information which is not obvious at first sight. Imagine a student who does her bachelor thesis in the field of Semantic Systems. Since she is not yet familiar with this topic she searches for people she could ask and relevant papers she could read. So she searches for the *topic* Semantic Systems and via the annotated link *belongs-to* she gets a list of *theses* that address this *topic* as well as a list of *material* that *covers* this *topic*. Since she prefers to be instructed by a colleague she asks the system for possible discussion candidates and gets all the information related to the queried person.

3.2 Scenario 2: Semantic Weblog

3.2.1 Problem

Consider students who document their learning experiences and progress in Weblogs. With growing information it becomes more and more difficult to find information not only for the students themselves but also for other people who might be interested in that documentation. Even though domain related information could be found by full-text search, certain types of information and how information is related could not be found that easily. There might be blog posts regarding literature reviews, video posts, posts with bibliographic information, posts discussing talks or events, and much more. In addition to defined blog post types, information might relate in one way or another even across blog posts or entire blogs. So a lecturer might be mentioned who is also author of publications a student has collected in a bibliographic overview. One of these publications might be discussed in a blog post and in another blog post the video talk referring to this publication might be published. Finally, the publication addresses a certain topic. In a usual weblog each of these information pieces

might be found sooner or later, but it would be difficult to find out about the relationships. Blog posts themselves would be found via search engines such as Technorati but only based on full text search. If one were interested only in posts including literature or event discussions he would not easily find them, since individual posts usually are not supplemented with metadata that could be used by search engines for indexing.

3.2.2 Solution

In order to solve this problem, a weblog could be semantically enriched. One aspect is to include a plug-in which helps to define the type of the post. Such plug-ins have already been developed[4]. They are based on the SIOC standards which helps to semantically connect online communities and their content. If a student would discuss an event such as a conference in a post she would select the appropriate post category. The post would then automatically be attached with according metadata. If another student would be interested in event discussions, because she is searching for a conference she could attend, she could search for a post type, connected to a topic, and would find all event discussions. Another possibility to support a student would be to support her in finding all the relevant posts from the blogosphere and thus colleagues who relate to a subject she herself discusses in a post. A plug-in like Yahoo!Shortcuts would automatically detect named entities such as persons, organisations, events, or locations within the posts of the student and automatically connect them to posts addressing the same object. When clicking on the highlighted object the student would find a lot of related and relevant information which helps her to get a better overview or deeper insight into the discussed subject. Of course, not only students benefit from semantically enhanced Weblogs. A teacher might want to improve her posts with e.g. illustrating pictures for providing rich learning content. She could be helped by e.g. the Open Calais plug-in Tagaroo[5] which extracts possible tags for the post and suggests Flickr[6] pictures based on the selected posts.

3.3 Scenario 3: Semantic Community Management

3.3.1 Problem

A common feature of resource management systems is that students and lecturers can create a profile. Such a profile usually contains personal data including e.g. affiliations or memberships. Also, it allows for managing resources such as publications, theses, courses, and even dates or websites. These profiles have the function to present oneself, to make oneself findable, and to organise personal resources. In these profiles, valuable information is contained, however rarely exploited for learning purposes. Despite all the information being available, the profiles do not support e.g. finding co-authors, members for a learning group, or experts for a joint proposal.

3.3.2 Solution

If profiles were enriched e.g. by areas of interest or expertise, and colleagues could be added as in Social Networking Systems, the profiles could be analyzed for similarities using Social Network Analysis. In a visualization students could search for colleagues who are near to a certain topic or who are near to her because of similar interests. Thereby points of contact for cooperation and collaboration could easily be identified. Facebook[7] offers such a service: If one has a favourite film he can query for all the people who favour the film, too, with her close friends being ranked before other people. Since students usually are members of various

[4] e.g. http://structuredblogging.org/
[5] http://www.opencalais.com/
[6] http://flickr.org/
[7] http://www.facebook.com/

Social Networking Services, amongst others StudiVZ[8], Facebook or MySpace[9], it would be helpful if the distributed contact information were integrated. Here the ideas of the SIOC initiative would help. On the basis of Community Ontologies, all the information could be exploited.

4 Conclusion and Outlook

As this article shows, Social Semantic Technologies integrate the advantages of both, Social Software and Semantic Technologies. They preserve the high flexibility of Social Software and bring in the semantics of Semantic Technologies without requiring strict formalizations that would keep people from using it. And it is not only a theoretical business far away from applicability. As the scenarios in the previous chapter show, there are various situations in the educational environment where Social Semantic Technologies support the development, retrieval, distribution, and acquisition of knowledge very effectively. Yet, we admit that the utilization of Social Semantic Technologies requires more effort and planning than the utilization of mere Social Software. So, in the case of a Semantic Wiki it must be first analysed for what exact situation it should be utilized. The same is true for Semantic Blogging. The detection of named entities such as locations or organizations might make no sense, since these are seldom subject to learning. The detection of topics might be much more helpful. However, the more careful planning has the positive effect that the task-technology fit is better analysed than it is when Social Software is applied. A simple Wiki or Weblog can that easily be implemented that thorough analyses might be relinquished.
Regarding the future, in our opinion Social Semantic Software will be the medium of choice when it comes to making user generated content better accessible and richer. As it was the case in the previous years with Social Software, further semantic plug-ins will be developed until there will be a consolidation to the most relevant ones. Some will remain a gimmick, some will be of real value. Next to Social Semantic Software, Semantic Technologies will further develop, since there are application areas which require a strict structure. So, librarian issues such as organizing print or digital books, articles, and magazines will be better solved by e.g. a pre-defined, rather stable domain ontology which allows the user to re-fined once searched articles.

Acknowledgement

The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

References

[1] Back, A., Gronau, N., and Tochtermann, K. (Hrsg.). (2008). Web 2.0 in der Unternehmenspraxis – Grundlagen, Fallstudien und Trends zum Einsatz von Social Software. Oldenbourg Wissenschaftsverlag, ISBN 978-3-486-58579-7.
[2] Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web. Scientific American, May 2001. Retrieved from http://www.sciam.com/article.cfm?id=the-semantic-web.
[3] Cayzer, S. (2006). What Next for Semantic Blogging. From Visions to Applications, In: Schaffert, S. und Sure, Y. (Eds.): Semantic Systems. From Visions to Applications. OCG Verlag, Wien (2006).

[8] http://www.studivz.net/
[9] http://www.myspace.com/

[4] Daconta, M. C., Obrst, L. J., and Smith, K. T. (2003). The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management, John Wiley & Sons.
[5] Duarte, F., Mattos, B., Bestavros, A., Almeida, V., and Almeida, J. (2007). Traffic Characteristics and Communications in Blogosphere. International Conference on Weblogs and Social Media, 2007, University of Maryland, Baltimore County, Maryland.
[6] Ebner, M., Zechner, J., and Holzinger, A. (2006). Why is Wikipedia so successful? Experiences in establishing the principles in Higher Education. Proceedings of the 6th International Conference on Knowledge Management I-KNOW 06, September 2006.
[7] Gantz J.F. (2007). The Expanding Digital Universe: A Forecast of Worldwide Information Growth Through 2010, IDC white paper, http://www.emc.com/about/destination/digital_universe/
[8] Gruber, T. (2006). Collective Knowledge Systems: Where the Social Web Meets the Semantic Web. 5th International Semantic Web Conference, November 7, 2006. Retrieved from http://tomgruber.org/writing/CollectiveKnowledgeSystems.pdf
[9] Nagler, W.; Saranti, A.; Ebner, M. (2008). Podcasting at TU Graz: How to Implement Podcasting as a Didactical Method for Teaching and Learning Purposes at a University of Technology; Proceeding of the 20th World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-Media), 2008, p.3858-3863
[10] O'Reilly, T. (2004). What is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. Retrieved from www.oreilly.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html.
[11] Pauschenwein, J., Jandl, M., Riegler, A., and Vasold, G. (2006). How to use weblogs in eSupervision? Proceedings of the 6th International Conference on Knowledge Management I-KNOW 06, September 2006.
[12] Pellegrini, T. and Blumauer, A. Was ist das Social Semantic Web? Magazin der Österreichischen Gesellschaft für Künstliche Intelligenz, 2007, 26, 19–23.
[13] Pellegrini, T. and Blumauer, A. (2008). Social Semantic Web – Die Konvergenz von Social Software und semantischen Technologien. In Back, A., Gronau, N., Tochtermann, K. (Hrsg.). (2008). Web 2.0 in der Unternehmenspraxis – Grundlagen, Fallstudien und Trends zum Einsatz von Social Software. Oldenbourg Wissenschaftsverlag, ISBN 978-3-486-58579-7.
[14] Schaffert, S. (2006). Semantic Social Software: Semantically enabled Social Software or Socially enabled Semantic Web. In S. Schaffert, Y. Sure (Eds.). Semantic Systems. From Visions to Applications. Wien: OCG Verlag.
[15] Schaffert, S., Bischof, D., Bürger, T.,Gruber, A., Hilzensauer, W., and Schaffert, S. (2006). Learning with Semantic Wikis. In M. Völkl (Ed.). First Workshop SemWiki2006 – From Wiki to Semantics, 2006, 109-123. Retrieved from http://www.eswc2006.org/technologies/usb/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf#page=117.

Authors:

Gisela, Granitzer, Dr.
Know-Center Graz, department
Inffeldgasse 21a, 8010 Graz
ggrani@know-center.at

Klaus, Tochtermann, Prof. Dr.
Know-Center Graz, Graz University of Technology, Joanneum Research
Inffeldgasse 21a, 8010 Graz
ktochter@know-center.at

Patrick, Hoefler
Know-Center Graz
Inffeldgasse 21a, 8010 Graz
phoefler@know-center.at