Semantics to the Bookmarks: A Review of Social Semantic Bookmarking Systems

Simone Braun
(FZI Research Center for Information Technologies
Simone.Braun@fzi.de)

Claudiu Schora
(FZI Research Center for Information Technologies
Claudiu.Schora@fzi.de)

Valentin Zacharias
(FZI Research Center for Information Technologies
Valentin.Zacharias@fzi.de)

Abstract: In this paper we present a review of systems that follow the novel paradigm of Social Semantic Bookmarking. Social semantic bookmarking allows for the annotation of resources with tags extended by semantic definitions and descriptions that also evolve (collaboratively) within the same system. We analyzed nine different systems that extend social bookmarking in the direction of more semantics; i.e. that enable their users to add semantics to the folksonomy. We studied the systems regarding the realization of the social semantic bookmarking paradigm, the features offered to the users to add semantics, what kind of semantics can be added, and how the system makes use of the semantics. We will present commonalities, main differences and distinctive features, and future trends.

Keywords: social bookmarking, social tagging, semantic tagging, social semantic bookmarking
Categories: H.3.4, H.5.3, L.1.3, L.1.4

1 Introduction

With the Web 2.0 many different kinds of novel, collaborative, and user centered applications have emerged. Aided by their ease of use, they have been remarkably successful and have found millions of users. Some of the most popular of these Web 2.0 applications are social tagging applications, in particular social bookmarking applications. These systems enable users to annotate web resources (e.g. bookmarks) with the arbitrary terms (so called tags) that they find most suitable for describing the resources. The popularity of these applications has shown that this principle of classification is much easier to use and understand then controlled vocabularies. These tags, however, are completely unstructured. Problems such as homonyms, synonyms, multilinguality, typos or different ways to write words, and tags on different levels of abstraction hamper their use for search and retrieval [Golder, 2006]; in particular in complex domains.

Recently, novel approaches have emerged that enable users to add semantics to the tags in order to overcome the limitations caused by the lack of semantic precision. These approaches allow extending tags with background knowledge such as synonymous names, super-, sub-topic relations or natural language descriptions. These approaches follow the paradigm of Social Semantic Bookmarking [Braun, 2008]. Social semantic bookmarking allows for the annotation of resources with tags extended by semantic definitions and descriptions that also evolve (collaboratively) within the same system. Similar to tagging approaches, new tags can be created whenever a need arises. Unlike these approaches, tags can have powerful descriptions and can be interlinked; for example allowing the system to understand that 'swimming bath' and 'swimming pool' are synonyms for the same concept. These powerful tag descriptions are similar to those used in traditional semantic annotation, but social semantic bookmarking allows for the simple adding and changing these descriptions within the same system and at the same time as the tags are used.

This paper presents a review of nine existing systems that follow the paradigm of social semantic bookmarking. The systems are analyzed and compared in their realization of the social semantic bookmarking paradigm, the features they provide to their users for adding semantics, the kind of semantics can be added, and how the system makes use of the semantics. We will present commonalities, main differences and distinctive features, and future trends.

2 Social Semantic Bookmarking Systems

We analyzed nine different systems: BibSonomy[1], SOBOLEO[2], Fuzzzy[3], GroupMe![4], Twine[5], ZigTag[6], Faviki[7], gnirz[8] and Annotea[9]. Only systems under active development that we could try ourselves (either with a publicly accessible system or through installation on our systems) are included – criteria that lead to the exclusion of the gnizr and the Annotea system at the end. All systems allow their users to extend tags used for annotating with additional semantics. However, each of the systems realizes social semantic bookmarking in a different way – providing their users different features to add additional semantics. Social semantic bookmarking is a relatively new area – BibSonomy launched at the beginning of 2006, the youngest, Faviki, at the beginning of 2008[10] – and the systems vary in technical maturity. Most

[1] http://www.bibsonomy.de
[2] http://tool.soboleo.com
[3] http://www.fuzzzy.com
[4] http://groupme.org
[5] http://twine.com
[6] http://zigtag.com
[7] http://faviki.com
[8] http://code.google.com/p/gnizr/
[9] http://www.annotea.org/
[10] Indeed the Annotea system is older but lacks the 'social' part of Social Semantic Bookmarking, e.g. a website where the bookmarks from other users can be discovered.

of the systems started as R&D projects, over the half originated in academia. The development of every system is still in progress.

2.1 Analyzed Systems

In the following we will shortly introduce each system in order to give a quick overview on the main characteristics of the analyzed systems:

- **BibSonomy:** BibSonomy [Hotho, 2006] is a system for the management of bookmarks of internet resources and publication entries. BibSonomy is a research project of the Knowledge and Data Engineering Group of the University of Kassel, Germany, that has launched the system at the beginning of 2006. BibSonomy offers functionality similar to that of well-known social bookmarking services but also functionality specifically tailored towards academics – e.g., sophisticated support for uploading and exporting bibliographic information in bibtex format. At its core, Bibsonomy differs from social bookmarking services by additionally offering users the possibility to create broader/narrower relations between tags. However, tag relationships are only local, i.e., each user can (and has to) maintain his own relationships and cannot profit from others' contributions in that respect.
- **SOBOLEO:** SOBOLEO [Zacharias, 2007] is a system for annotating and organizing bookmarks of internet resources with concepts from a shared ontology. SOBOLEO is a research project developed since the beginning of 2007 at the FZI Research Center for Information Technologies Karlsruhe, Germany. Within SOBOLEO every user owns and shares his bookmarks with the whole community, i.e. a bookmark is added one time to the repository and there is one annotation that every user can edit. SOBOLEO's main unique functionality is that it integrates social bookmarking with ontology development; i.e. the users of one community collaboratively maintain and develop the ontology they use for the annotation.
- **Fuzzzy:** Fuzzzy [Lachica, 2008] is a system for managing bookmarks of web pages and ISBN numbers. Fuzzzy is developed within the PhD project of Roy Lachica at the University of Oslo and its development started at the end of 2006. It is based on Topic Maps technology. Besides hierarchical and related tag relations, the users can choose of 22 specific predefined association types to link tags. These tag relations apply to the whole system and are editable by other users. Another main concept is voting for gardening and maintenance: the users can vote on bookmarks, tags a bookmark is annotated with, relations between tags, and users.
- **GroupMe!:** GroupMe [Abel, 2007] attempts to bridge the gap between the Semantic Web and Web2.0 with an RDF based social bookmarking application. The Semantic Web Group at the University of Hannover in Germany has been developing GroupMe! since 2007. The main unique functionality of GroupMe! is the extension of the tagging idea with the concept of 'groups' (collections): all annotated internet resources (websites, music, videos, photos, and news feeds) can be organized into groups. These form another level of information that can be used for browsing and search.
- **Twine:** is a system that supports organizing and sharing bookmarks of web pages, images, videos, products, and books. Twine is a commercial product run by Radar Networks has been publicly usable since October 2007. The main organizing

principle of Twine is so called 'twines', a kind of user interest group. The users can join or create such interest groups in order to share bookmarks. Through bookmarks grouped in the same twine group, tags are set into relation. Another functionality of Twine is the faceted filtering of search results; i.e. the users can filter by seven categories like tags, people, places, or item type.

- **ZigTag:** ZigTag allows tagging and organizing bookmarks of internet resources. ZigTag is run by ZigTag Inc., a small Canadian company, and had its beta release in April 2008. In contrast to the other approaches, the users can annotate their bookmarks with predefined tags, i.e. they choose from a given list of tags with a specific meaning. The users can additionally add their own tags for private usage. They cannot add further semantics. Tags are set into relation by the system.
- **Faviki:** The Faviki, launched in May 2008, is developed and by the Serbian Web designer Vuc Milicic. The system distinguishes itself by relying on Wikipedia terms for the annotation of bookmarks, i.e. users annotate with tags that are titles of Wikipedia articles. The Faviki system does not allow to use tags not contained in Wikipedia or to add semantic relations between tags.
- **ginzr:** ginzr is a system for managing bookmarks of internet resources using tags and folders. ginzr provides the users the possibility to define broader/narrower, related, and member-of relationships between tags according to the SKOS vocabulary [Miles, 2009]. Additionally, the users can add geospatial information to the bookmarks. The system has been published as open source application under Mozilla Public License 1.1 by the Image Matters LLC company in March 2008. There is no publicly available installation. The system can be downloaded and hosted on one's own. However, despite several attempts, it was not possible to get the system to work properly. Therefore, we excluded this system from the overall comparison as we could not evaluate the system provider's information.
- **Annotea:** Annotea [Koivunen, 2006] is a metadata standard for semantic web annotations, it is implemented in a number of tagging tools and server applications. Annotea and its implementations have been developed by the W3C. Annotea differs from other approaches to social tagging in its emphasis on standards on decentrality, that it has sharing of bookmarks among services build in from ground up. However, Annotea has not been under active development for some years and for this reason was excluded from the overall comparison.

3 Analysis

To give a comprehensive overview of the respective strength and weaknesses of the systems introduced above, Table 1 details the main discriminating features among the applications. The features used for the comparison relate to four main aspects: tags, relations, additional bookmark information, and miscellaneous functionalities. They will be discussed in detail in the following.

	Bib-Sonomy	SOBO-LEO	Fuzzzy	Group-Me!	Twine	ZigTag	Faviki
Tags							
freely selectable / editable	x / x	x / x	x / x	x / x	x / x	x / x	x / -
meaning definable	-	x	x	-	-	-	-
editing others' tags	-	x	x	-	-	-	-
multilingual / automatic translation	- / -	x / -	x / -	- / -	- / -	- / -	x / x
auto completion	x	x	x	-	-	x	x
tag suggestion	x	-	x	-	x	x	x
rating of tags	-	-	x	-	-	x	-
Relations							
system internal	x	-	-	x	x	x	x
user defined	x	x	x	x	-	-	-
type definable	-	x	x	-	-	-	-
multiple / transitive	x / -	x / x	x / -	x / -	- / -	- / -	- / -
public / local	- / x	x / -	x / -	x / -	- / -	- / -	- / -
rating of relations	-	-	x	-	-	-	-
Add. Bookmark Information							
title suggestion	x	x	x	x	x	x	x
content suggestion	x	-	x	x	x	-	-
add selected text	x	-	-	-	x	-	x
title / content analysis	x / x	- / x	- / x	- / -	x / x	x / -	x / x
rating of bookmark	x	-	x	-	x	x	-
Misc. Features							
bookmarklet tool	x	x	x	x	x	x	x
bookmark import	x	-	-	-	x	x	-
bookmark export	x	x	x	x	-	x	-
RSS / REST API	x / x	x / x	x / x	x / x	x / -	x / -	x / -
discussion / interactive help	- / -	x / -	x / -	- / -	x / x	x / -	- / -
user incentives	-	-	x	-	-	-	-
activity logging	x	x	x	x	x	x	-
guest account	-	x	-	x	-	-	-
open source	-	-	-	-	-	-	-

Table 1: Comparison of Social Semantic Bookmarking Systems

3.1 Tags

Within all systems the users can freely select which and how many tags they want to use for annotating bookmarks. All but one system allow the editing and creation of tags; only the Faviki system restricts users to select tags from a predefined list of Wikipedia terms; i.e. only tags that have a reference to Wikipedia articles can be used for annotation. If users want to use other tags, they first have to create a new Wikipedia article. Users of the other systems can freely edit the tags they want to use. New tags are automatically added to the system when the annotation is saved. However, GroupMe! users cannot modify once added tags, i.e. errors made during annotation cannot be corrected.

SOBOLEO and Fuzzzy offer their users the possibility to specify the meaning of a tag by adding additional information like a description or synonymous terms. The specification of synonymous terms can also be used for basic multilingualism support. However, Faviki is the only system that provides real multilingualism and automatic translation support for tags. Within ZigTag the provider predefine the specific meaning of tags, supporting in this way the disambiguation of syntactically identical tags. The users can discuss the predefined tags but it is not obvious how this global list of predefined tags and their descriptions is generated. Self-created tags cannot be further specified. Within the remaining systems, tags are limited to the string itself. [Abel, 2008] indicated for the GroupMe! system the usage of the property moat:meaning of the MOAT ontology [Passant, 2008] in order to specify a tag's meaning, however it seems to be not yet available in the current public system release. Changes to tags not declared as private apply to the whole system in SOBOLEO and Fuzzzy, in all other systems the information added to tags is private. In ZigTag the users can request new tags to be added to the global list of predefined tags.

During the annotation process, all systems except Twine and GroupMe!, support the users with auto completion for tags. BibSonomy, Fuzzzy, Twine, ZigTag, and Faviki further provide tag suggestions. Fuzzzy users additionally can rate if a tag that is applied to a bookmark is suitable or not, whereas ZigTag users can discuss a tag in general.

3.2 Relations

All analyzed systems provide support for relations between tags: set by the system itself, defined by the users or both. BibSonomy allows both – it sets internally 'related' and 'similar' relations between tags by computing co-occurrence for the first and cosine similarity for the latter. These relations are used to support tag-based navigation. The users can also define broader/narrower relations between two tags, e.g. 'pet' -> 'dog'. He can relate one tag with multiple other tags. However, defining types of relations other than broader/narrower is not possible. Furthermore, these relations are only locally valid; i.e. for two tags of one user. Searching for resources annotated with a specific tag, e.g. 'pet', within one user's repository also retrieves resources only annotated with narrower tags, e.g. 'dog'. Only narrower relations are considered for retrieval. It is not possible to find resources only annotated with 'pet' when searching for 'dog'. Likewise, transitivity is not considered as well; i.e. if there

is also the relation ‘dog’ -> ‘poodle’, resources only annotated with ‘poodle’ cannot be found when searching for ‘pet’.

SOBOLEO and Fuzzzy leave the definition of relations to their users. In both systems it is possible to set multiple relations between tags. SOBOLEO users can define broader/narrower and related relations between two tags – according to the SKOS vocabulary. Fuzzzy users can add hierarchical and related relations which have to be further semantically specified by choosing one of 22 available association types. In both systems the relations apply systems wide. Likewise, the hierarchical and related relations support tag-based navigation in both systems. The specific association types within Fuzzzy are not further considered. Additionally, Fuzzzy users can rate the relations between tags. SOBOLEO’s search engine furthermore takes into account existing relations; i.e. it retrieves resources annotated with ‘poodle’ when searching for ‘dog’. It also provides query refinements and relaxations based on broader tags.

Within the GroupMe! system, users can define collections, so called ‘groups’ to which they can add resources or other groups. This results in the automatic creation of relations between the tags with which the resources within the same group are annotated. These relations are used for search support, e.g. searching for tag ‘a’ also retrieves resources only annotated with tag ‘b’ if both tags are used within the same group. Thus the relations are globally valid. Even though groups can be added to other groups, the relations are not considered as transitive. Other types of relations are not possible.

Twine, ZigTag, and Faviki only provide support for relations set by the system. Faviki relies on DBpedia and the semantic relations it provides, which is mainly the support for multiple languages. Within the ZigTag system there are predefined relations between tags with synonymous meaning, e.g. ‘NYC’ and ‘New York’, and relations for related tags (called “tags in common with”), however, it is not obvious how these relations are generated (e.g. for the latter case co-occurrence seems to be used but not every co-occurring tag is listed). This is also true for Twine. Twine extracts via content analysis if a tag represents people, places, organizations, and types of items. And it can also identify tags that are related. This information is used to support filtering of research results.

3.3 Additional Bookmark Information

All presented systems provide suggestions for the bookmark title during the annotation process. Half of the systems, BiboSonomy, Fuzzzy, GroupMe!, and Twine, also make suggestions to describe the content of the bookmark. Within BibSonomy, Twine, and Faviki the user can select a paragraph within the web page to be annotated and add this via drag’n’drop to the content description. The content description in GroupMe! is not editable by the user. All systems also analyze the title or content of the resource, e.g. for tag suggestions (BibSonomy, Twine, ZigTag, Faviki), semantics extraction (Twine), or full text indexing (SOBOLEO, Fuzzzy, Twine; facilitating search in the first two systems).

The users of the Fuzzzy system can rate a bookmark as positive or negative; Twine and ZigTags users can leave comments on a bookmark and BibSonomy users can decline a bookmark as spam.

3.4 Miscellaneous Features

As additional feature all of the systems provide a bookmarklet that can be installed in the browser and to allow users to quickly tag and store a bookmark for the currently opened internet resource from within the browser.

Users can import existing bookmarks into BibSonomy, Twine, and ZigTag. BibSonomy allows the import of bookmarks from delicious or of publications via an EndNote and BibTex files. ZigTag users can import their delicious or browser bookmarks; Twine additionally supports the import from digg.

Regarding export, there is a wide range of diverse functionalities and formats – ranging from one user's bookmarks annotated with a specific tag to the whole repository dump. Each system provides some export functionality via RSS feeds. BibSonomy provides various additional export formats for bookmarks and publications, such as XML, BURST, SWRC, BibTex, and EndNote. However, the defined semantic relations between tags cannot be exported. SOBOLEO can export the whole bookmark repository together with the ontology or only the ontology in the SKOS format. Likewise, Fuzzzy provides on request an export of all bookmarks and the ontology in XTM format. GroupMe! offers an RDF based export functionality using FOAF, DCMI, Tag ontology and the proprietary GroupMe! ontology as schema. ZigTag users can export their bookmarks in the Netscape Bookmark Format; similar to BibSonomy, any semantic relations are lost in the export. BibSonomy, SOBOLEO, Fuzzzy, and GroupMe! additionally provide a web service API using Representational State Transfer (REST), e.g. to post a new bookmark. Twine and Faviki don't offer any API or export functionality beyond RSS feeds.

Due to RSS feeds the users of any system can stay up-to-date. BibSonomy displays the most recently added bookmarks and mostly used tags on the frontpage. It additionally shows the three most popular bookmarks over the last 7, 30, or 120 days. SOBOLEO displays the newest bookmarks on top within the browse area. Fuzzzy differentiates between newly added tags and bookmarks and bookmarks annotated with one of the user's favorite tags. GroupMe! users can see the latest activities within their groups as 'Dashboard News'. Similar functionality is offered by Twine for a user's twine group and by ZigTag respectively.

Discussions are supported by half of the systems. SOBOLEO integrates a chat for synchroneos communication into the ontology editor. Fuzzzy users can make propositions for improvement on which other users can comment and vote. ZigTag and Twine both have an integrated forum. All systems offer instruction guidelines or tutorial on how to use the system, but only Twine provides an interactive help directly integrated with the systems UI elements.

The Fuzzzy system highly encourages its users to engage in community activities. They can make propositions for improvement, there are special voluntary tasks for gardening, users can be voted to get special privileges or they can gain so called 'karma points' as incentive for each contribution. It is the only system offering such functionality.

None of the system is open source. Each system requires user registration in order to use the systems. SOBOLEO and GroupMe! additionally offer guest accounts to test the systems without registering.

4 Conclusions

As a general conclusion, there is a big interest to extend social bookmarking in the direction of more semantics. Social Semantic Bookmarking allows a group of users to collaboratively create and evolve an index of resources together with the powerful semantic vocabulary used to organize it. Social Semantic Bookmarking promises better retrieval, better use of annotation, better integration of the repository with semantic web infrastructure etc. while avoiding the problems commonly associated with semantic annotation approaches – such as a high initial cost to build ontologies.

The analyzed systems show a wide range of how social semantic bookmarking can be realized and supported. One major distinctive aspect is the degree of freedom of what users can contribute and what is done automatically. Fuzzzy and SOBOLEO highly rely on the user community. Because both systems aim at supporting the collaborative development of ontologies, they give their users the greatest latitude to freely define and edit tags and relations. BibSononomy and GroupMe! limit their users to one specific type of relations they can add. The commercial systems ZigTag and Twine aim in offering highly automated support. The semantics in these systems are internally derived and predefined, the users have only little influence. At the other end Faviki relies on standardization based on Wikipedia/DBpedia, only allowing the users to select tags for annotation.

Another characteristic is the scope and impact of the (user added) semantics. Within BibSonomy user added relations are only locally relevant and set on top of one user's tags. Within SOBOLEO and Fuzzzy changing tags and semantics have a system wide effect on every user. The additional semantics are mainly used to support better retrieval and tag-based navigation: e.g. on the retrieval of all bookmarked pages for one tag, the bookmarks associated with its subtags are also returned. All kinds of relations are used to facilitate navigation from one tag to another. The SOBOLEO system also uses the semantic information to augment the full text search of the contents of the bookmarked internet resources. Formal semantics seem not to play an important role with the current generation of systems; to the authors best knowledge none of the systems employs an actual inference engine.

There is a big interest in particular to tackle the problem of how tagging data can be exchanged between systems. All the systems are offering (often multiple) ways to export the created annotations. However, at the same time this analysis shows that there still is considerable disagreement about the most important features and – even more crucially – about suitable formats to exchange the tagging data. Without an agreement in this domain, the promise of exchanging tagging data can obviously not be achieved.

Social Semantic Bookmarking applications hold a huge potential for future development as part of the developments towards a Web 3.0 as a user-centered semantic web. However, in addition to the technical questions about the right functionality of a Social Semantic Bookmarking system there is also a lack of understanding of the emergence and evolution of semantics as part of everyday collaborative activities. The research community still lacks a comprehensive model that describes the process of such a collaborative knowledge maturing; a model that then could also guide the development of support mechanisms. The authors will

further explore this with their SOBOLEO system as part of the European Integrating Project MATURE[11].

References

[Abel, 2007] Abel, F., Henze, F.M., Krause, D., Plappert & D. Siehndel, P.: Group Me! Where Semantic Web meets Web 2.0. In: Proc. of the 6th Int. Semantic Web Conf., 2007

[Abel, 2008] Abel, F.: The benefit of additional semantics in folksonomy systems. In Proc. of the 2nd PhD Workshop on information and Knowledge Management (Napa Valley, California, USA, October 30 - 30, 2008). CIKM '08. ACM, New York, NY, 2008, pp. 49-56.

[Braun, 2008] Braun, S., Zacharias, V., Happel, H.-J.: Social Semantic Bookmarking. In Proc. of 7th International Conference on Practical Aspects of Knowledge Management - PAKM'08 (Yokohama, Japan, November 21-23, 2008), LNCS 5345, Springer Verlag Berlin Heidelberg, 2008, pp. 62-73

[Golder, 2006] Golder, S., Huberman, B.A.: The Structure of Collaborative Tagging Systems. J. of Inf. Sc. Vol. 32(2), 2006, pp. 198-208

[Hotho, 2006] Hotho, A., Jäschke, R., Schmitz, C.; Stumme, G.: BibSonomy: A Social Bookmark and Publication Sharing System. In: CS-TIW'06. Aalborg: Aalborg University Press, 2006

[Koivunen, 2006] Koivunen, M.-R.: Semantic Authoring By Tagging with Annotea Social Bookmarks and Topics. In: Proceedings of the Semantic Authoring and Annotation Workshop at the International Semantic Web Conference ISWC'06, 2006

[Lachica, 2008] Lachica, R., Karabeg, D.: Quality, Relevance and Importance in Information Retrieval with Fuzzy Semantic Networks. In: Proc. of TMRA 2008, University of Leipzig LIV series post-proceeding Paper, 2008

[Miles, 2009] Miles, A., Bechhofer, S.: SKOS Simple Knowledge Organization System Reference. W3C Candidate Recommendation 17 March 2009, 2009

[Passant, 2008] Passant, A., Laublet, P.: Meaning Of A Tag: A collaborative approach to bridge the gap between tagging and Linked Data. In Proceedings of the WWW 2008 Workshop Linked Data on the Web (LDOW2008), Beijing, China, 2008.

[Zacharias, 2007] Zacharias, V., Braun, S.: SOBOLEO – Social Bookmarking and Lightweight Ontology Engineering. In: Proc. of the Workshop on Social and Collaborative Construction of Structured Knowledge at WWW'07, CEUR Workshop Pro. Vol. 273, 2007

[11] http://mature-ip.eu