

Alexandre Sigel, Lars Bünthen, and Marcus Frank: Topic Maps-based Semblogging with semblog-tm.
In: L. Maicher, A. Sigel, and L.M. Garshol (Eds.): TMRA 2006, LNAI 4438, pp. 227-242, 2007
© Springer-Verlag Berlin Heidelberg 2007
This is a legal self-archived preprint of the author-created version. For the published version, see
<http://www.springer.com/dal/home/computer/computer+series?SGWID=1-40099-69-173621840-0>

Topic Maps-based Semblogging with semblog-tm

Alexander Sigel, Lars Bünthen, and Marcus Frank

University of Cologne, Department of Information Systems and Information Management,
Pohligstr. 1, 50969 Köln, Germany
sigel@wim.uni-koeln.de

Abstract. Semantic blogging combines blogs with the Semantic Web for improved metadata. Our analysis of six semblogging approaches and systems reveals that all are RDF-based, suffer from not using Published Subjects as proxies for subjects, and do not employ semantic relations motivated by knowledge organization. In contrast, we introduce and discuss semblog-tm, a prototype of a Topic Maps- and PSI-based semblogging system with a well-motivated set of semantic relations. Its basic requirements have been derived from the three main use cases: managing lightweight ontologies, attaching semantics to blog entries, and providing aggregated semblogging data as semantic knowledge services to other semantically-enabled systems. The four main system components (blogging, topic map, PSI, semantic knowledge services), and its capabilities are sketched. The web application semblog-tm is implemented as a plugin for the blojsom blogging engine, using Java and Tomcat with Velocity templates. Topic map operations are realized via TMAPI with TM4J or OKS, and Axis is employed for web services.

Keywords: semantic blogging; semblogging; social software; published subjects; PSIs; emergent lightweight ontologies; distributed knowledge management; knowledge services

1. Introduction

The emerging field of semblogging (for short SB) combines blogs with the Semantic Web to enrich blog entries with more explicit and machine-understandable metadata, relating both to structure and content. This semantics is conveyed by semantic tags (semtags), connected by semantic associations.

Blogging [1, 2], in particular knowledge blogging in distributed/decentralized knowledge management (DKM) [3] aids personal and collaborative knowledge work [4, 5]. With this instrument, knowledge workers can easily record and publish rather short notes (blog entries) on subjects they want to discourse about. Blogs and wikis are prominent exponents of social software [6, 7], a paradigm in which effects within human groups emerge, and groups are primary objects within the system. Mechanisms like blogrolls, trackbacks or pingbacks, and tagging can lead to highly inter-networked and interactive communities focused around subjects.

To improve findability and syndication of blog entries, knowledge workers associate short text strings (tags) with their blog entries. Those tags are either reused or newly created. A blogger can use or create as tag any string (s)he wants to use as a retrieval cue. In general, a tag describes an aspect of the aboutness of an entry, but it can be any pragmatic aspect a user likes to attribute. Collaboratively tagging blog entries is an extreme form of user-oriented, social, or democratic indexing [8]. It leads to folksonomies (which can be visualized as tag clouds), and hence to folksonomic tagging [9-13] as a kind of thematic annotation of blog entries.

1.1 Relevance and Motivation: Why We Need Blogging with More Explicit Structure and Semantics

“Blogging, as a subset of the web as a whole, can benefit greatly from the addition of semantic metadata. The result — which we will call Semantic Blogging — provides improved capabilities with respect to search, connectivity and browsing compared to current blogging technology.” [14]

In conventional blogging with tags, structure and semantics are not explicit. According to [14], “(...) blogs, and the posts that they contain, lack sufficient semantic information regarding the topics that they are talking about or how the current topic under discussion relates to previous blog discussion threads”. This shortcoming makes syndication of blog entries from different blogs about the same subject difficult, since computers have to guess what is only implicit. Improved blogging systems should thus provide support for both structural and content-related metadata.

Structural metadata identifies and describes elements internal to one blog entry, the relations between those elements, or the form of related blog entries [14]. Structured blogging [15, 16] allows bloggers publishing of structured information such as reviews or events by use of machine-interpretable templates. *Content-related (semantic) metadata* refers to the meaning of tags and relations. It expresses the subjects the entries are about. At present, tags are typically only strings used as names, not connected with a concept. It is not explicitly and publicly described or defined which concept a tag shall denote and be a computer proxy for. Except evolutionary competition, there is no vocabulary control of and no central registry for such tags. This may be regarded as an advantage of social software [17], but may also lead to chaos. Another problem is that tags are “flat”, i.e. no typed semantic relations are specified between tags, although bloggers might want to express not only tags, but semantically interrelated tags.

A more explicit semantic markup or annotation of content like blog entries with semantic web means is needed such that semantic web applications can automatically process the content [18]. Some knowledge workers want to express and use knowledge structures (make assertions) while blogging, be it for personal or joint knowledge work. They are longing for tools supporting them in expressing more than just text with simple tags.

A major problem is how to aggregate and virtually collocate blog entries about the same subject of thought, and how to systematically present related blog entries. A simple example of this type of problem are book reviews scattered across blogs [19]. Not only are the entries physically distributed all over the blogosphere, but independent bloggers annotate from their personal perspective, not adhering to a central ontology. Social bookmarking, or folksonomic tagging [7] suffers – more than traditional

indexing in knowledge organization – from different yet acceptable viewpoints and conceptualizations, from problems with polysemy and homonymy, quality issues, and from rather arbitrary tags.

1.2 The Semblogging Approach

SB [20] crosses blogs with the semantic web, extending blogging by more explicit semantics. “[B]logs [are] enriched with semantic, machine-understandable metadata” [21]. A few system conceptualizations and implementations exist, see e.g. [14, 18, 20]. SB is a special case of semantic annotation [22-24] in line with distributed knowledge management. Blog contents become a data source for semantic web applications and services and can be aggregated and repurposed on a semantic basis in a smarter way.

Some terminology: [25] calls semantic tags *semtags* and the process of semantic tagging *semtagging*. In analogy we will call the process of blogging with such tags semantic blogging (*semblogging*), and such bloggers *sembloggers*. Occasionally, though not consistently used, *to semtag*, used as verb, stands for associating a single semtag with a blog entry, while *to semblog*, used as a verb, stands for semantic blogging in a more general sense (several acts of semtagging). Although not only semtags in isolation, but semantic relations (“*semrels*”) are used to form complex statements, with semtags playing roles, for the sake of simplicity we will subsume this under semtagging/SB.

The challenge of SB is how to best support bloggers in making key structure and semantics explicit while retaining the ease of blogging. Most needed are lightweight means such that bloggers can optionally define the semantics of their folksonomic tags and relate them to each other with typed semantic relations. Such SB systems can be realized with a variety of semantic technologies, in particular with RDF/OWL vs. Topic Maps.

1.3 Research Gap at Project Start in 2004

Cayzer’s seminal SB work has inspired several SB systems, but there is no comparative discussion, and no such system is freely available as open source. There is no work on interoperably interlinking several SB systems such that aggregated assertions can be exchanged and queried. Semantic knowledge services between such systems are only emerging. Published Subjects are not used in SB. No reusable tools and no public registry service can be identified for PSIs [26]. Relation types are often ad-hoc based. Often, rather simple links are employed instead of links with role types. No one has replicated RDF SB work with Topic Maps to better understand the issues in SB systems. Based on our intuition (or bias?) that Topic Maps have presumed strengths, in particular for the expression of knowledge structures by humans, and on our observation that those strengths are not fully explored and exploited for augmenting SB [26], we wonder how a Topic Maps-based SB system would look like.

1.4. Our Contribution

Within the *kpeer* (knowledge peers) project (2004–2006) we investigated Topic Maps-based SB with particular interest in the exchange of assertions in this setting. Three diploma students significantly contributed to this research, investigating the conceptual design and implementation of the initial system [27], its extension by aspects of distribution and semantic knowledge services [28], and interoperable semantic knowledge services with Topic Maps and RDF, with *semblog-tm* as a data provider [29]. To have better control, we decided to build our own system.

Our analysis of six SB approaches and systems revealed that all are RDF-based, suffer from not using Published Subjects as proxies for subjects, and do not employ semantic relations motivated by knowledge organization. In contrast, we introduce and discuss *semblog-tm*, a fully implemented prototype of a Topic Maps- and PSI-based SB system with a well-motivated set of semantic relations. It uniquely combines the features: Topic Maps [30–32] as an interesting alternative to RDF(S)/OWL, Published Subjects [33–36], and Semantic Relations explicitly based on work in knowledge organization [37–38].

We have shown that a Topic Maps-based SB system is both feasible and useful. Compared to other SB systems, some added-value is offered, warranting further work. Our main contributions are:

- Identification and short characterization of six RDF-based SB conceptualizations and systems
- Conceptualization and prototypical implementation of the first public and open-source Topic Maps- and PSI-based SB system, and of the only SB system based on association types motivated by knowledge organization theory.
- The first SB system with open semantic knowledge services, aggregating the data and providing them back to any topic map fragment consumers. Implementation and initial discussion of Topic Maps-based semantic knowledge services, advancing the interoperability and scalability of SB systems
- Establishing a stronger argument for Topic Maps-based SB, and for integrating appropriate information architecture and semantic interoperability into other SB systems.

2. Review of Related Work on Semblogging

2.1 RDF-based Semblogging

The Seminal Semblogging Demonstrator (HP)

Within the European research project SWAD-E (Semantic Web Advanced Development for Europe), Steve Cayzer and colleagues at the Hewlett-Packard laboratories in Bristol coined the term semantic blogging and developed the very first prototypical SB demonstrator [20, 39]. The special application scenario was collaborative bibliography management in a scientific community (SB for bibliographies). The demonstra-

tor can be tried out online¹, and there is a user guide [40]. Requirements [41] and lessons learnt [42] are documented as SWAD deliverables. This prototype supports better semantic view, navigation and query. In principle, arbitrary vocabularies can be used, e.g. such represented with SKOS [43-45]. The system is implemented using the Java blogging platform blojsom² and Jena³ for RDF.

A Semblogging Client within Haystack (MIT)

Karger and Quan built a SB client into Haystack⁴ (“a platform for authoring end user semantic web applications”) incorporating publishing, aggregation, and browsing capabilities. The scenario assumes a scientist attending a conference who wants to know more about the talks presented at the conference. The conference program is available in RDF and can be viewed with Haystack as a semantic web browser. For each abstract, corresponding blog entries discussing the paper are suggested. The relations between blog entries (here: the structure of the discussion) can be graphically depicted. After viewing all blog entries (containing both approving and disagreeing contributions), the scientist can create a blog entry annotating the most convincing blog entry or argument. As with the HP system, blojsom is used.

Semblog - Personal Knowledge Publishing Suite

Ohmukai, Takeda, Numa and colleagues proposed a personal knowledge publishing suite, co-incidentally called “semblog”, to support the information exchange between individuals [46-48]. This SB system extends RSS aggregators with two capabilities:

1. users can categorize and republish RSS feeds, and
2. standalone aggregators (for each user one), called glucose, form a p2p network.

A user can recommend feeds read to other users. Links between bloggers are modelled with FOAF. Three methods of egocentric search help finding related entries: entries linked from the annotated resource, entries annotating the same resource, and entries classified with the same keyword.

Semantic Blogging Research Project (KMI)

In the SB research project⁵ at Open University, Bertrand Sereno, Marc Eisenstadt, and Simon Buckingham Shum investigate how blogs can be used for distributed knowledge management and construction [49]. SB is seen as a special case of distributed sense-making and argumentation. The knowledge mapping tool “Compendium”⁶ helps better structuring and interrelating of blog entries, extending the entries with semantics. Discussions spanning several blogs can be marked up with semantic relations and displayed in structured form.

¹ <http://www.semanticblogging.org/>

² <http://wiki.blojsom.com/wiki/display/blojsom/About+blojsom>

³ <http://jena.sourceforge.net/>

⁴ <http://haystack.lcs.mit.edu/>

⁵ <http://kmi.open.ac.uk/projects/semanticblog/>

⁶ <http://www.compendiuminstitute.org/>

Aggregating Entries with Language Games in the Tagsocratic Project

Avesani, Cova, Hayes, and Massa [19] describe “problems in aligning similar concepts created by a set of distributed, autonomous users”. They claim that, employing their tagsocratic approach, entries about the same topic can be aggregated even if bloggers have used tags with different names. For a given tag in one blog, a (for the moment) central mapping service returns tags from different blogs. As a common reference structure, so-called “global tags” form a kind of meta- or switching thesaurus.

semiBlog – Semblogging to Publish Desktop Data

semiBlog applies the concept of SB to publish existing desktop data with semantic markup to better integrate desktop data in the communication process [14, 21, 50]. As a proof-of-concept, semiBlog⁷, a prototypical SB editor and blog reader has been implemented as a desktop application. High emphasis is put on usability and tight integration with desktop applications. An example scenario assumes annotating a scientific paper after discussing it with a colleague. The blogger can annotate the blog by drag-and-drop-referencing information objects available on the semantic desktop, like his colleague (selecting from his address book), the project (selecting the URL of the project page) and the bibliographic reference (selecting the reference in the bibliographic tool of his choice). Other users can import the annotations from the blog to their desktop applications. SB is thus a way to easily exchange semantic web data. In contrast to the HP demonstrator and the SB client within Haystack which use blojsom, semiBlog uses the WordPress blog engine. The SIOC (Semantically-Interlinked Online Communities) ontology developed by the same authors is employed in a SIOC plugin for WordPress. The recent semiBlog is currently the most advanced RDF SB system and unique in its combination of structural and content metadata (see [14], fig. 5: A combined SIOC and semiBlog graph). Compared to Cayzer’s HP demonstrator, some semantic relations are used, but they are not yet very rich.

2.2 Topic Maps-based Semblogging

At least to Topic Mappers, Topic Maps seem an obvious choice for SB. However, despite all our efforts, we could not identify scientific references or published systems on SB realized with Topic Maps (except our own earlier ideas [26, 55-57]. In 2004, Jack Park had presented SB as one example for Augmented Storytelling [58]. Dmitry Bogachev had informally discussed SB with topic maps, but we have no further details. Lars Marius Garshol had realized a simple SB application using Ontopia’s OKS, which he kindly provided to us. In his blog entry on “[t]ags/folksonomies and Topic Maps”⁸, he also shared some thoughts why “[t]opic Maps are the perfect way to fill the gap between what tagging is and what it should be.” Interest in Topic Maps-based SB is slowly rising, see e.g. [59].

⁷ <http://semiblog.semanticweb.org/>

⁸ <http://www.garshol.priv.no/blog/33.html>

2.3 Discussion

There seems no lack of motivation for SB, but a lack of adaption in blogging communities. Until recently, this was partially because there were no tools available in public, but we expected the first usable SB systems to become public in the near future. Except in the semiBlog project, usability issues have been neglected. It should be easier to use SB systems, and they should be tighter integrated into the work environment and knowledge processes of knowledge workers.

To the best of our knowledge, the very few existing SB systems are based on RDF, not on Topic Maps. For example, Karger and Quan, in stating “[w]e characterize the notion of semantic blogging – the publication of RDF-encoded Web logs“ [18], do not even seem to be aware of Topic Maps as an alternative for encoding semblog meta-information. No discussion can be found on using Topic Maps vs. RDF(S)/OWL for SB. RDF and Topic Maps can be made semantically interoperable in some way [60-65], and to a certain extent this is also true for RDF(S) and OWL, e.g. using TMCL. Compared to RDF alone, Topic Maps seem better suited for humans to express knowledge, and converting “down” from Topic Maps to RDF is losing information (context and identity) [66]. This suggests that it is worthwhile to explore basing SB directly on Topic Maps, at least as an alternative.

Although PSIs are useful to ground semantics both with RDF and Topic Maps, all the systems suffer from the Web’s “identity crisis” [34], because subject indicators [33, 35, 36] are not used, In addition, some systems lack semantic relationships between tags, or the relations are rather simple.

3. semblog-tm, a Topic Maps-based Semblogging System

We shortly report on semblog-tm as demonstrated at TMRA 06. A demonstrator was available online.⁹ Because a demonstration video, together with a transcript [67] and slides [68] is available, screenshots are not included here. The source code with a system installation guide is available via anonymous SVN sourceforge checkout.¹⁰

3.1 Conceptual Design and Main Capabilities

With the solution elements Topic Maps, Published Subjects and a predefined set of semantic relations already in mind, we developed several scenarios [28, 29], decomposed them into main use cases, and derived basic requirements. Starting from a rather centralized approach, we later explored aspects of distribution, and provided semantic knowledge services both internally and externally. We grouped into three main use cases:

1. Managing lightweight, decentrally governed ontologies: Defining ontologies (sem-tags connected by semantic associations)

⁹ <http://semlblog.wim.uni-koeln.de/blojsom/blog/> (no longer supported)

¹⁰ <http://semlblog.sf.net/>

2. SB: Attaching semantics to blog entries by connecting semblog entries with entries from those ontologies, and displaying them
3. Managing aggregated SB data as semantic knowledge services: Provide other semantically-enabled systems with services on topic map fragments.

System capabilities as seen from the user's perspective include:

Semtag: discover, display/follow, copy third-party semtag to own ontology space, attach semtag to blog entry, define own semtag from scratch, mint PSI for semtag, associate two semtags. *Semantic retrieval*: various predefined queries, like: all semtags by all sembloggers, or blog entries for specified semtags, same-server or cross-server retrieval, result display, export to .xtm.

3.2 System Architecture, Tool Selection, and Implementation

We designed semblog-tm to have the following four high-level system components and suggest this as a possible architecture for other systems (cf. fig. 1):

1. A *blogging component* dealing with all conventional blogging functions.
2. A *topic map component* for manipulating topic maps. It provides capabilities for topic map manipulation, merging and querying, importing, exporting, and format conversion. For each semblogger, a separate topic map is held.
3. A *PSI component* allows the creation of new PSIs, thereby aiding in the unambiguous identification of concepts.
4. A *semantic knowledge services component*. Web services manipulating topic map fragments. Internally, these realize a service-oriented architecture, externally, they ease synchronization, aggregation and querying of assertions between SB servers, or a SB server and other semantic applications.

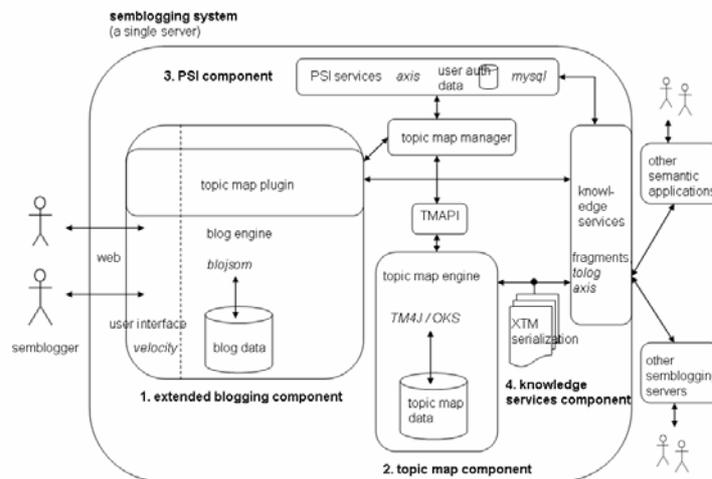


Fig. 1. semblog-tm main components of the system architecture: Simplified view

The core topic map is designed to support both formal (structural) and semantic (content-related) metadata:

Formal metadata: Blog entries are redundantly held in blojsom and in the topic maps. A blog entry is connected via authorship (source – author) with a blogger. The topics *human-aboutness* and *machine-metadata* have been defined to type internal occurrences of Published Subjects, according to the OASIS Published Subjects TC recommendations¹¹ [33], and the recent proposal by Naito on metadata for Published Subjects [69]. In addition, a topic *primary-topic-subject-identifier* has been defined to discern the designated PSI from additional subject identifiers.

Content-related metadata: This encompasses the semtags themselves, their interrelations, and the relation aboutness between semtag: issue and source: blog entry. Based on a proposal for semantically differentiated thesaurus relations in knowledge organization [37, 38]¹², the SB system supports two class-1 relations and all 23 class-2 relations. An example for such a class-2 relation is: “TMRA06” (instrument) is claimed to be instrumental for “Leveraging the semantics” (purpose).

The web application semblog-tm is implemented as a plugin for the blojsom blogging engine, using Java and Tomcat with Velocity¹³ templates. Topic map operations are realized via TMAPI¹⁴ with TM4J or OKS. Open source software has been used as far as possible, and portability between Windows and Unix has been achieved. Blojsom was chosen for the *blogging component*, in addition to other advantages, because it is also used for the HP and the Haystack demonstrator. This should make comparison of and interchange between such SB applications easier. The *Topic Map component* is implemented on top of TMAPI, and TM4J (open source) and OKS (commercial) are supported topic map engines. We chose the SOAP web service framework Axis for the *knowledge services component*.

4. Discussion and Suggestions for Further Work

Usability and Users. Maturity: One reviewer found fault with not taking real (empirical) end user requirements into account, which results in additional burden for users, without evaluating and analyzing advantages and disadvantages of the proposed approach. Our primary goal was not provision of a system excelling in usability, albeit desirable. Rather did we pursue design science to better understand how such a system might look like. Thus we first creatively synthesized and imagined potential usage situations. Unfortunately, no user communities could be involved later. Before

¹¹ Recommendation 1 (“A Published Subject Indicator should provide human-readable metadata about itself”), and recommendation 2 (“A Published Subject Indicator may provide machine-processable metadata about itself”)

¹² In collaboration with Winfried Schmitz-Esser, PSIs for these associations have been defined and published, integrating the PSIs once defined by the company Seruba (now defunct).

http://web.archive.org/web/*/http://seruba.com

¹³ <http://jakarta.apache.org/velocity/>

¹⁴ <http://www.tmapi.org/>

wider use, several known usability issues would have to be dealt with. We suggest improving the system together with selected users, and evaluating it in a larger, thematically focused community.

Semantic Retrieval vs. Fulltext: It was also recommended to discuss why semantic retrieval with semtags and along their interrelations should outperform searching without additional tags (fulltext). The retrieval results much depend on the context. Although requiring too much metadata can be a heavy, even counterproductive burden, there is sufficient reason to believe fine-grained semantic retrieval has advantages, because precision and recall can be controlled in new ways. We have not evaluated our system with user communities, so future work should look into empirical retrieval tests. Work on SB agents parsing conventional blog entries, heuristically interpreting them, and republishing the adapted contents with semantic markup¹⁵ does not render SB meaningless. This solution cannot fully detect most structures which knowledge workers could explicitly express right from the beginning.

Knowledge Organization: One reviewer missed a more detailed justification of why we chose an approach for semantic relations based on knowledge organization, and why this particular one. Any approach based on one or more sound theories of semantic relations (e.g. knowledge organization, rhetorical structure theory, or didactic theory) would be better than ad-hoc relations. The Schmitz-Esser set of relations was chosen because the first author has long-standing experience in knowledge organization and is involved in work on transferring this particular relation set to Topic Maps-based applications. We do not claim that this particular arsenal of relations is the most useful or complete. However, this set turned out as quite powerful, although additional sets such as rhetorical ones (like agreeing/disagreeing) should be integrated in further work. By design, only semtags can be copied to own ontology views, not semrels (associations). Further improvements should support sembloggers in adding new relation types and subtypes, not only instances, and in changing the particular fixed (but configurable) set of predefined association types to a more flexible approach. For ease of use, relation types could be coupled with templates.

RDF(S)/OWL: One reviewer criticized why we value RDF-based approaches without detailed reasons. Another correctly pointed out that the graph model RDF cannot be blamed for the fact that some RDF-based prototypical blogging software does not support semantic relations between blog entries and semtags, as this could be added with RDFS/OWL, both well integrated into RDF. Therefore, we should not base our argument for using Topic Maps on dismissing RDF. In this paper, we were not at all interested in arguing pro or contra RDF(S)/OWL vs. Topic Maps in general, but in drawing from Topic Maps for improved SB. Our main motivation was design science-oriented. We do not claim that Topic Maps are in general superior to other approaches, or that certain Topic Maps-based SB features could not be achieved with other means. In fact, both Published Subjects and well-founded semantic relations can be integrated into other semantic technologies, and should so. However, we do claim that our combination of features:

- was lacking in or not the focus of previous work,

¹⁵ See e.g. ongoing diploma thesis “Entwurf und Realisierung von Semantic Blogging Agents” at Technical University Munich, <http://www11.in.tum.de/lehre/da/proj-776.phtml>, last accessed 2006-06-12

- is useful for SB systems in general, independent of their technical basis, and
- can be well realized with Topic Maps in a more natural way than with RDF alone, or with RDFS/OWL.

We did neither intend nor achieve an in-depth review and feature comparison of all the SB systems. We are aware that some RDF-based systems partially exhibit more sophisticated features, or can be used more easily. In particular, a closer look and intertwining with the recent semiBlog system is desirable. In general, the RDF(S)/OWL and Topic Maps communities should still come closer together, better connecting research, learning from each other, and fostering convergence and interoperability. Future work should better amalgamate advantages of both approaches.

Semwikis, e.g. ArtificialMemory: Related to SB, but not treated here, are semantic wikis (semwikis) [6, 10, 51, 52]. After finishing this paper, we became aware of Artificial Memory¹⁶, a SemWiki developed by Lars Ludwig [53, 54]. Facts treated as subdocuments can be expressed and linked with semantic relations, also in blog style.

Blogging Component: Future work might adapt our blojsom plugin to other blogging engines.

PSI Component: We only implemented it because a reusable third-party solution did not exist. Future work should integrate more flexible and sophisticated third-party services.

TMAPI and Topic Map Engines: Basing all topic map operations on TMAPI by design has also several major drawbacks: Only the TMAPI layer operations can be used, which are slower and more cumbersome than directly using the more powerful capabilities of engine-specific native framework APIs. This forces some duplication in navigation and querying capabilities. Tolog via TMAPI is not yet supported, so we had to awkwardly access topics with the TMAPI index system, and for complex queries, the engines had to be queried directly. Because the tolog version in TM4J was not sufficient for our purposes, semantic retrieval in semblog-tm currently depends on OKS. Future work should implement higher-level operations on top of TMAPI. Support for more backends using TMAPI, and for different topic map engines not using TMAPI is also desirable. semblog-tm should also be changed from in-memory to database backend, although the topic maps are frequently serialized as .xtm.

Semantic Knowledge Services: We had to decide for one of the different approaches for topic maps services. As shown with TMSHare (TopicMapster) [70], topic map fragments can be generated and p2p-like shared. Some alternatives are the TMRAP protocol [71-73], including capabilities for tolog queries, or TMIP [74]. We decided not to use TMRAP, although very useful, because TMAPI does not yet support it. Except Axis, we could not look into other service frameworks, like e.g. WSMO tools¹⁷. Better service interoperability as tackled by a recent harmonization initiative would be nice, and so would be more open semantic knowledge services.

For the **future of Topic Maps-based SB research**, we foresee more research on:

- p2p-like distribution for all components, e.g. TMRAP and semantic overlay in p2p networks with topic map data, or distributed search over PSIs.

¹⁶ <http://www.artificialmemory.net/>

¹⁷ <http://www.wsmo.org/>

- federated interoperability of PSI and topic map aggregation services, as well as more work on interoperability between Topic Maps and RDF(S)/OWL, between topic map query languages and Sparql, and work on semantic knowledge services.
- experiments with PSI registries with additional metadata for Published Subjects
- trust and reputation building for semtags
- empirical research with and actual use in SB communities
- open distributed knowledge management applications with topic maps in general.

References

1. Walker, J.: Weblog. In: Herman, D.H., Jahn, M., Ryan, M.-L. (eds.): The Routledge Encyclopedia of Narrative Theory. Routledge, London and New York (2005) 45
2. Picot, A., Fischer, T. (eds.): Weblogs professionell. Grundlagen, Konzepte und Praxis im unternehmerischen Umfeld. d.punkt (2005)
3. Staab, S., Stuckenschmidt, H. (eds.): Semantic Web and Peer-to-Peer. Decentralized Management and Exchange of Knowledge and Information. Springer, Berlin (2006)
4. Röhl, M.: Knowledge Blogs: Persönliche Weblogs im Intranet als Werkzeuge im Wissensmanagement. In: Picot, A., Fischer, T. (eds.): Weblogs professionell. Grundlagen, Konzepte und Praxis im unternehmerischen Umfeld. d.punkt (2005) 95-111
5. Kleske, J.: Wissensarbeit mit Social Software. Konzeption & Entwicklung eines Systems für die kollaborative Wissensarbeit in der Forschung basierend auf Social Software. Diplomarbeit Fachhochschule Darmstadt, Darmstadt (2006)
6. Schuster, M., Rappold, D.: Social Semantic Software – was soziale Dynamik im Semantic Web auslöst In: Pellegrini, T., Blumauer, A. (eds.): Semantic Web: Wege zur vernetzten Wissensgesellschaft. Springer (2006) 189-199
7. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social Bookmarking Tools (I) – A General Review. D-Lib Magazine 11(4) (2005) <http://www.dlib.org/dlib/april05/hammond/04hammond.html>
8. Sigel, A.: How can user-oriented depth analysis be constructively guided? In: Beghtol, C., Howarth, L.C., Williamson, N.J. (eds.): Dynamism and Stability in Knowledge Organization. Procs. 6th Int. ISKO Conference. 10-13 July 2000, Toronto, Canada. Ergon, Würzburg (2000) 247-253
9. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. Journal of Information Science 32(2) (2006) 198-208
10. Voß, J.: Collaborative thesaurus tagging the Wikipedia way (2006) <http://arxiv.org/abs/cs/0604036>
11. Mathes, A.: Folksonomies – Cooperative Classification and Communication Through Shared Metadata. Computer Mediated Communication – LIS590CMC, Graduate School of Library and Information Science, University of Illinois Urbana-Champaign (2004) <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf>
12. Quintarelli, E.: Folksonomies: power to the people. ISKO Italy Conference, 2005-06-24, Mailand, Italy (2005) <http://www.iskoi.org/doc/folksonomies.htm>
13. Folksonomy (2006-06-08) <http://en.wikipedia.org/wiki/Folksonomy>
14. Möller, K., Bojars, U., Breslin, J.G.: Using Semantics to Enhance the Blogging Experience. In: Sure, Y., Domingue, J. (eds.): The Semantic Web: Research and Applications, 3rd European Semantic Web Conference, ESWC 2006, Budva, Montenegro, June 11-14, 2006, Proceedings, LNCS 4011. Springer, Berlin (2006) 679-696
15. Structured blogging (2006-06-08) <http://www.structuredblogging.org/>

16. Paquet, S.: Towards structured blogging [Blog entry] (2003-03-13) <http://radio.weblogs.com/0110772/stories/2003/03/13/towardsStructuredBlogging.html>
17. Shirkey, C.: Ontology is Overrated: Categories, Links, and Tags (2005) http://shirky.com/writings/ontology_ouerrated.html
18. Karger, D.R., Quan, D.: What would it mean to blog on the semantic web? *Journal of Web Semantics* 3(2) (2005) 147-157
19. Avesani, P., Cova, M., Hayes, C., Massa, P.: Learning Contextualised Weblog Topics. WWW2005, 2nd Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, 2005-05-10, Chiba, Japan (2005) <http://sra.itc.it/people/cova/data/papers/paper-www-workshop-05.pdf>
20. Cayzer, S.: Semantic Blogging and Decentralized Knowledge Management. *Communications of the ACM* 47(12) (2004) 47-52
21. Möller, K., Breslin, J.G.B., Decker, S.: semiBlog - Semantic Publishing of Desktop Data. 14th Conference on Information Systems Development (ISD2005), August 2005 Karlstad, Sweden (2005) <http://sw.deri.org/~knud/papers/semiBlogISD2005Word/semiBlogISD2005print.pdf>
22. Handschuh, S., Staab, S.: *Annotation for the Semantic Web*. IOS Press, Amsterdam (2003)
23. Reif, G.: Semantische Annotation. In: Pellegrini, T., Blumauer, A. (eds.): *Semantic Web: Wege zur vernetzten Wissensgesellschaft*. Springer (2006) 405-418
24. Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E., Ciravegna, F.: Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Journal of Web Semantics* 4(1) (2006) 14-28
25. Dill, S., Eiron, N., Gibson, D., Gruhl, D., Guha, R., Jhingran, A., Kanungo, T., Rajagopalan, S., Tomkins, A., Tomlin, J.A., Zien, J.Y.: SemTag and Seeker: Bootstrapping the Semantic Web via Automated Semantic Annotation. WWW 2003 (12th Int. Conf. on World Wide Web). ACM, Budapest, Hungary (2003) 178-186
26. Sigel, A.: Report on the Open Space Sessions. In: Maicher, L., Park, J. (eds.): *Charting the Topic Maps Research and Applications Landscape*. First International Workshop on Topic Map Research and Applications, TMRA 2005, Leipzig, Germany, October 6-7, 2005, Revised Selected Papers, LNAI 3873. Springer (2006) 271-280 http://www.wim.uni-koeln.de/uploads/media/Report_from_the_open_space_sessions_v24_with_Copyright.pdf
27. Frank, M.: *Semantic Blogging mit Topic Maps – Entwurf und Implementierung eines Prototypen*. Diplomarbeit. Wirtschaftsinformatik, insbesondere Informationsmanagement. Universität zu Köln, Köln (2005)
28. Bünten, L.: *Verteiltheit und semantische Wissensdienste für Topic Maps-basiertes Semblogging: Konzeption und prototypische Implementierung*. Diplomarbeit. Wirtschaftsinformatik, insbesondere Informationsmanagement. Universität zu Köln, Köln (2006)
29. Jansen, S.: *Interoperable semantische Wissensdienste mit Topic Maps und RDF: Konzeption und prototypische Implementierung*. Diplomarbeit. Wirtschaftsinformatik, insbesondere Informationsmanagement. Universität zu Köln, Köln (2006)
30. Biezunski, M., Bryan, M., Newcomb, S.R.: *ISO/IEC 13250 Topic Maps (Second Edition). Approved revised text (2002)* <http://www.jtc1sc34.org/repository/0322.htm>
31. Garshol, L.M., Moore, G.: *Topic Maps — Data Model*. International Standard (for review). (2006-06-18) <http://www.isotopicmaps.org/sam/sam-model/>
32. Garshol, L.M., Moore, G.: *Topic Maps — XML Syntax*. Final Draft International Standard Draft (for review). ISO 13250: Topic Maps (2006-06-19) <http://www.isotopicmaps.org/sam/sam-xtm/>
33. Pepper, S.: *Published Subjects: Introduction and Basic Requirements*. OASIS Published Subjects Technical Committee Recommendation (2003-06-24) <http://www.oasis-open.org/committees/download.php/3050/pubsubj-pt1-1.02-cs.pdf>

34. Pepper, S., Schwab, S.: Curing the Web's Identity Crisis: Subject Indicators for RDF. XML Conference 2003, Philadelphia, Pennsylvania, USA (2003) http://www.idealliance.org/papers/dx_xml03/html/abstract/05-01-05.html
35. Maicher, L.: Subject Identification in Topic Maps in Theory and Practice. In: Tolksdorf, R., Eckstein, R. (eds.): Berliner XML Tage 2004. XML-Clearinghouse, Berlin, Germany (2004) 301-307
<http://www.informatik.uni-leipzig.de/~maicher/publications/%5Bmaic04cFINAL%5D.pdf>
36. Pepper, S.: The Case for Published Subjects. WWW2006 Workshop: Identity, Reference, and the Web (IRW2006), Edinburgh, Scotland (2006) http://www.ontopia.net/topicmaps/materials/The_Case_for_Published_Subjects.pdf
37. Schmitz-Esser, W.: Thesaurus and Beyond: An Advanced Formula for Linguistics Engineering and Information Retrieval. Knowledge Organization 26(1) (1999) 10-22
38. Introduction to Integrative Cross-Language Ontology (ICLO): Formalizing and interrelating textual knowledge to enable intelligent action and knowledge sharing. In: Schmitz-Esser, W. & Sigel, A. (2006): Introducing Terminology-based Ontologies. <http://eprints.rclis.org/archive/00006612/>, pp. 54-113. Ninth International ISKO Conference 2006
39. Cayzer, S.: Semantic Blogging: Spreading the Semantic Web Meme. XML Europe 2004, 18-21 April, Amsterdam, NL (2004) http://idealliance.org/papers/dx_xmle04/papers/03-05-03/03-05-03.pdf
40. Cayzer, S.: Semblogging - a guide to the HP Labs semantic blogging demonstrator (2003) <http://www.semanticblogging.org/semblog/whatisit.html>
41. Cayzer, S., Shabajee, P.: SWAD-Europe deliverable 12.1.2: Semantic Blogging and Bibliographies – Requirements Specification (2003) <http://www.w3.org/2001/sw/Europe/reports/pdf/12.1.2.pdf>
42. Reynolds, D., Cayzer, S., Shabajee, P., Steer, D.: SWAD-Europe deliverable 12.1.8: SWAD-E Demonstrators – Lessons Learnt (2004) <http://www.w3.org/2001/SW/Europe/reports/demo-lessons-report/>
43. Miles, A., Brickley, D.: SKOS Core Vocabulary Specification. W3C Working Draft 2 November 2005 (2005) <http://www.w3.org/TR/swbp-skos-core-spec>
44. Miles, A., Brickley, D.: SKOS Core Guide. W3C Working Draft 2 November 2005 (2005) <http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/>
45. Miles, A., Matthews, B., Beckett, D., Brickley, D., Wilson, M., Rogers, N.: SKOS: A language to describe simple knowledge structures for the web. XTech 2005: XML, the Web and beyond. Idealliance, Amsterdam, NL (2005) <http://www.idealliance.org/proceedings/xtech05/papers/03-04-01/>
46. Ohmukai, I., Numa, K., Takeda, H.: Egocentric Search Method for Authoring Support in Semantic Weblog. Workshop on Knowledge Markup and Semantic Annotation (Semantot2003), collocated with the Second International Conference on Knowledge Capture (K-CAP2003), 2003. (2003)
47. Ohmukai, I., Takeda, H., Numa, K., Hamasaki, M., Adachi, S.: Personal Knowledge Publishing Suite with Weblog. 13th International World Wide Web Conference (WWW2004), Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, 2004-05-18. New York, NY, USA (2004) http://www.kasm.nii.ac.jp/papers/takeda/03/kcap03_ohmukai.pdf
48. Ohmukai, I., Takeda, H.: Semblog: Personal Publishing Platform with RSS and FOAF. In: Brickley, D., Decker, S., Guha, R.V., Miller, L. (eds.): Proceedings of the 1st Workshop on Friend of a Friend, Social Networking and the (Semantic) Web, September 2004, Galway, IE (2005) 217-221 http://www.w3.org/2001/sw/Europe/events/foaf-galway/papers/pp/semblog_personal_publishing_platform/
49. Sereno, B.: Knowledge Mapping as a blogging approach [Blog entry] (2005) <http://kmi.open.ac.uk/people/bertrand/index.php/2005/07/15/knowledgemapping-as-a-blogging-approach/>

50. Möller, K., Decker, S.: Harvesting Desktop Data for Semantic Blogging. In: Decker, S., Park, J., Quan, D., Sauermann, L. (eds.): Proceedings of 1st Workshop on The Semantic Desktop - Next Generation Information Management & Collaboration Infrastructure. ISWC 2005, November 6, 2005 (CEUR Workshop Proceedings), Galway, Ireland (2005) 79-91 <http://sw.deri.org/~knud/papers/semiBlogSemDesk2005/semiBlogSemDesk2005.pdf>
51. Völkel, M., Krötzsch, M., Vrandečić, D., Haller, H., Studer, R.: Semantic Wikipedia. 15th International Conference on World Wide Web, WWW 2006, May 23-26, 2006, Edinburgh, Scotland (2006)
52. Völkel, M., Schaffert, S.: Proceedings of the First Workshop on Semantic Wikis - From Wiki to Semantics (SemWiki2006), co-located with the 3rd Annual European Semantic Web Conference (ESWC 2006). Budva, Montenegro, June 12, 2006 (2006) <http://www.aifb.uni-karlsruhe.de/WBS/hha/papers/SemanticWikipedia.pdf>
53. Ludwig, L., O'Sullivan, D., Zhou, X.: Artificial Memory Prototype for Personal Semantic Subdocument Knowledge Management (PS-KM). Demonstration Proposal. ISWC 2004 (2004) <http://iswc2004.semanticweb.org/demos/36/paper.pdf>
54. Zhou, X., Li, Q., Ludwig, L., Chen, Y.: Subject-Oriented Knowledge Formalization: Method and Prototype. IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceeding. IEEE Computer Society, 18-22 December 2006, Hong Kong, China (2006) 853-858
55. Sigel, A.: Organisation verteilter Wissens mit semantischen Wissensnetzen und der Aggregation semantischer Wissensdienste am Beispiel Digitale Bibliotheken/Kulturelles Erbe. In: Ohly, H.P., Sieglerschmidt, J., Swertz, C. (eds.): Wissensorganisation und Verantwortung. Gesellschaftliche, ökonomische und technische Aspekte. Ergon, Würzburg (2006) 276-292 http://www.wim.uni-koeln.de/uploads/media/ISKO2004_SIGEL_Organisation_verteilten_Wissens_v16_01.pdf
56. Sigel, A.: Content Intelligence durch Verknüpfung semantischer Wissensdienste am Beispiel von Semblogging. 2. Kongress Semantic Web und Wissenstechnologien. Semantic Web Services, Darmstadt, Germany (2005) http://www.wim.uni-koeln.de/uploads/media/Content_Intelligence_ZGDV05_v16.pdf
57. Sigel, A.: kPeer (Knowledge Peers): Informationssuche beim verteilten SemBloggen. Workshop "P2P Information Retrieval in Deutschland", Leipzig (2005) <http://www.sempir.informatik.uni-leipzig.de/fohlen/sigel.pdf>
58. Park, J.: Augmented Storytelling on the Web. National Storytelling Conference, July 2004, Bellingham, WA, USA (2004) <http://www.nexist.org/nsc2004/index.html>
59. Rajbhandari, S., Andres, F., Naito, M., Wuwongse, V.: Semantic-augmented support in Spatial-Temporal Multimedia Blog Management. In: Procs. Of TMRA 2006, this volume (2007).
60. Garshol, L.M.: Living with Topic Maps and RDF: Topic Maps, RDF, DAML, OIL, OWL, TMCL. XML Europe Conference 2003, London, England (2003) http://www.idealliance.org/papers/dx_xmle03/index/title/da78f2d953553c194e0c0dbed6.html
61. Garshol, L.M., Naito, M.: RDF and Topic Maps Interoperability in Practice / Realization of seamless knowledge: connecting distributed RDF and Topic Maps. 3rd International Semantic Web Conference (ISWC2004) / SIG-SWO, Hiroshima, Japan (2004) <http://iswc2004.semanticweb.org/demos/19/paper.pdf>
62. Ontopia: The RTM RDF to Topic Maps Mapping: Definition and Introduction. [Ontopia technical report]. Version 0.2. Ontopia (2003) <http://www.ontopia.net/topicmaps/materials/rdf2tm.html>
63. Pepper, S., Presutti, V., Garshol, L.M., Vitali, F.: Guidelines for RDF/Topic Maps Interoperability. W3C Editor's Draft 30 June 2006 (2006) <http://www.w3.org/2001/sw/BestPractices/RDFTM/guidelines-20060630.html>

64. Pepper, S., Vitali, F., Garshol, L.M., Gessa, N., Presutti, V.: A Survey of RDF/Topic Maps Interoperability Proposals. W3C Working Group Note 10 February 2006 (2006) <http://www.w3.org/TR/rdfm-survey/>
65. Pepper, S., Vitali, F., Garshol, L.M., Gessa, N., Presutti, V.: Approaches to RDF/Topic Maps Interoperability. XTech 2005: XML, the Web and beyond, Amsterdam, NL (2005) <http://www.idealliance.org/proceedings/xtech05/papers/03-03-01/>
66. Garshol, L.M.: Q: A model for topic maps: Unifying RDF and topic maps. Extreme Markup Languages 2005 (2005) <http://www.mulberrytech.com/Extreme/Proceedings/html/2005/Garshol01/EML2005Garshol01.html>
67. Sigel, A., Bünten, L.: Topic Maps-based Semblogging with semblog-tm: System demonstration at TMRA06. Video and edited transcript (2006) <http://www.wim.uni-koeln.de/uploads/media/semblog-tm-demo-at-TMRA06.zip>
68. Sigel, A., Bünten, L.: Topic Maps-based Semblogging with semblog-tm: System demonstration at TMRA06. Slides (2006) http://www.wim.uni-koeln.de/uploads/media/TMRA_Semblogging_Demo_v15_2006-10-11.pdf
69. Naito, M.: Metadata for Published Subjects. Proposal for a new work item (ISO/IEC JTC 1/SC 34) (2006) Private communication
70. Ahmed, K.: TMSHare - Topic Map Fragment Exchange in a Peer-To-Peer Application. XML Europe Conference 2003, London, England (2003) http://www.idealliance.org/papers/dx_xmle03/papers/02-03-03/02-03-03.html
71. Garshol, L.M.: TMRAP - Topic Maps Remote Access Protocol. In: Maicher, L., Park, J. (eds.): Charting the Topic Maps Research and Applications Landscape, First International Workshop on Topic Maps Research and Applications, TMRA 2005, Leipzig, Germany, October 6-7, 2005, Revised Selected Papers, LNAI 3873. Springer (2006) 53-68 <http://www.garshol.priv.no/download/text/tmrappdf>
72. Pepper, S.: Towards Seamless Knowledge. Integrating Public Sector Portals. XML 2004 Conference, Washington, D.C., USA (2004) <http://www.ontopia.net/topicmaps/materials/Towards%20Seamless%20Knowledge.ppt>
73. Pepper, S., Garshol, L.M.: Seamless Knowledge. Spontaneous Knowledge Federation using TMRAP. Extreme Markup 2004 Conference, Montréal, Canada (2004) <http://www.ontopia.net/topicmaps/materials/Seamless%20Knowledge%20with%20TMRAP.ppt>
74. Barta, R.: TMIP, a RESTful Topic Maps Interaction Protocol. Extreme Markup Languages Conference 2005. IDEAlliance, Montréal, Canada (2005) <http://www.mulberrytech.com/Extreme/Proceedings/xslfo-pdf/2005/Barta01/EML2005Barta01.pdf>