Enterprise Wikis: Technical Challenges and Opportunities

Benedikt Kämpgen, Basil Ell, Elena Simperl, Denny Vrandečić, Frank Dengler
firstname.lastname@kit.edu

Abstract: Social software has proven valuable in enterprises for collaborative knowledge management. In order to introduce a wiki in the enterprise, we propose a solution that combines Web 2.0 and Semantic Web technologies. We describe how this solution resolves the technical challenges, beyond that, opens up new opportunities, and, also, how it can be realized in a concrete enterprise scenario.

1 Introduction

Social software as a tool for knowledge sharing and collaboration is gaining more and more relevance in the enterprise world [DRBM09]. This especially is true for so-called enterprise wikis, that, just as wikis in the public web, provide their advantages of low usage-barriers and direct benefits within a company intranet. However, simple provision of a Wikipedia-alike does not guarantee acceptance by employees; such wiki software needs to be customized to the specificities of the corporate context. There are technical, social, and organizational challenges to this customization. Human behavior and organizational habits impeding the adoption of social software have been discussed much recently [HDW10]. For instance, a collaboration-unfriendly corporate culture and an unclear value proposition for stakeholders hinder the adoption of Web 2.0 technologies.

Much less is known about the technical challenges related to this adoption. Integration into the enterprise IT landscape and compliance with diverse internally and externally defined policies and regulations are the most obvious examples. Existing literature deals with technical challenges on a high-level without providing useful guidance for enterprises. Also, there are profit-driven systems exclusively fitted to the requirements of enterprises, however, how they solve technical challenges is not known to the public domain. Therefore, we want to elaborate on technical challenges of introducing an enterprise wiki. To do so we build upon our experiences earned in introducing wikis at three globally operating companies in the sectors of telecommunication, consultancy, and electronics design, namely British Telecommunications plc, Accenture, and Cadence Design Systems. These experiences align along the theme of proposal development in response to customer demands, a scenario which is not only highly critical across many business sectors, but also can be seen as a representative example in which *enterprise knowledge structures* are collaboratively created, enriched, and exploited. Together with potential users, we have identified requirements that impose technical challenges to a wiki-based solution.

Ankolekar et al. already have argued that Web 2.0 and the Semantic Web provide complementary technologies [AKTV07]. Therefore, we assume that applying semantic concepts

to an enterprise wiki will not only help to overcome the technical challenges, but also provide new opportunities. More concretely, we propose to use as basis Semantic MediaWiki [KVV+07]. Its usage of standard semantic technologies such as RDF[1] and ontologies provides advantages beginning from an integrated means to formally describe the meaning and organization of the content to various enhancements of the way information is retrieved, displayed and navigated within the wiki. For evaluation, we implemented the solution; the results were again presented to potential users, who have confirmed that the requirements were met.

With this paper, it is our aim to achieve greater awareness of the technically motivated challenges behind enterprise wiki adoption and to allow enterprises to make an informed decision about deployment of our solution in a similar scenario. Developed around representative enterprise knowledge structures – information sources for proposal development –, we assume the technical challenges to be typical for enterprise wikis. The description of the implementation is not only meant for evaluation but also provides a concrete example.

The remainder of this paper is structured as follows. We first introduce the enterprise scenario and describe its relevant requirements (Section 2). In Section 3, we describe how semantic technologies can be used to fulfil the technical challenges. In Section 4 we foster our claim, explaining a concrete implementation. After that, an overview of related work is given in Section 5, followed by conclusions in Section 6.

2 Enterprise scenario

Proposal development in enterprises is commonly perceived as a knowledge-intensive, collaborative process, in which a proposal manager and a team create a description of the products and services delivered by the company at an estimated cost to a potential customer. A proposal includes various types of information – for instance, about marketing, pricing and certification – provided by various enterprise departments – for instance, technical consultants, product specialists and sales persons. The development typically includes activities such as *selecting the proposal team*, *gathering information about the customer*, *discussing customer issues and possible solutions*, and *getting approval for pricing*; still, it is highly variable and its full particulars can hardly be recorded through productivity software which is often used in this context, such as Microsoft Word, Excel and SharePoint, as well as messaging services [STW+ar]. To illustrate this, consider the activity of gathering information about a customer. The way this activity is carried out depends on the preferences and expertise of the proposal development team, and on undocumented social communication and collaboration practices. One might visit a website, consult the intranet portal or call a former colleague, to name just a few.

For engineering of functional and non-functional requirements, about 50 potential users – knowledge workers involved in proposal development – were asked within the ACTIVE project[2] for their opinion. We only include requirements that impose technical challenges

[1] http://www.w3.org/RDF/

[2] http://www.active-project.eu/

to our solution. Similarly, we do not include requirements that by themselves can be solved by typical wiki solutions, e.g., click-to-edit functionality, storage of both text and media, and change tracking.

Functional requirements A proposal development workspace should help users to record, share and collaboratively refine on relevant *enterprise knowledge structures*. Those structures comprise various information sources, available to an enterprise, that its employees use to manage the knowledge for their daily work. Examples include customer descriptions, product specifications, and price lists.

The system should provide guidance to the user about what information is to be put into the wiki. Also, users prefer to not depend on wiki syntax, but to have simpler and faster ways of adding information into the wiki, e.g., forms.

Users should be able to access the results of common activities executed for previous proposals and find such reusable pieces of content in due time.

The wiki is supposed to not only provide the necessary structure, but also to offer concrete information from already existing data sources relevant for proposal development. Enterprises contain many different data sources, for instance relational databases, content management systems, and various document formats. They contain more or less unstructured information. Examples include descriptions of finished proposals; price lists of competitors; reports about industry sectors; and other elements that employees use for proposal development and therefore want to access, discuss, and refine through the wiki.

Data quality is of high importance. Users will not adopt the wiki if incorrect information is contained and not distinguishable from relevant information. On the one hand, flexibility of what to put into the wiki should be preserved. On the other hand, it should be possible to discover and solve data problems.

Employees are accustomed to tools that help with developing proposals. For instance, the end-proposal is usually delivered in form of a Microsoft Word document. In the ideal case, users are free to continue using their tools, but these tools are extended and allow to exploit the wiki's added values. Also, as interaction in a wiki occurs asynchronously, some users prefer to be able to keep track of changes without constantly visiting the wiki.

Non-functional requirements The acceptance is likely to be higher if the tool is intuitive to use, also by users without technical background, and minimally invasive to established workflows and the enterprise IT landscape. The system should run sufficiently fast, with loading times similar to external webpages.

3 Solution based on a semantic wiki

In this section, we argue that semantic technologies will help to overcome much of the technical challenges imposed by our proposal development scenario, and beyond that, open up new opportunities. As a reference semantic wiki, we have chosen to use Semantic

MediaWiki (SMW). Then, in the next section, we present an implementation of a proposal development wiki having these capabilities.

Creating structured information Our proposal development scenario makes it necessary to capture and refine enterprise knowledge structures within the wiki. For that, guidance to the users is necessary. Information stored in SMW conforms to machine-readable RDF. More understandable, it allows to have property-value pairs explicitly assigned to wiki pages. Such a property-value pair can be a named link (so-called *object property*) to another page, e.g. "locatedInCountry" Page of country, but can also be a typed attribute (so-called *datatype property*), e.g., "hasTag" String, "hasFoundingDate" Date, and "hasHeight" Number. *Properties* can be inserted into a page with wiki syntax, but also using forms. First, this makes them easily usable. Second, enterprise knowledge structures, as we have them in our proposal development scenario, can be defined through *categories* of pages (so-called *classes*) with certain properties and serve as guidance for the users of how to use the wiki and what structured information to capture in it, e.g., proposal with a team, a customer, and a due date.

The structure can first be modelled in standard knowledge representation languages for the Semantic Web, such as RDF, and OWL[3]. Depending on the expressivity, this *ontology* can be automatically or manually applied to the wiki [VK06]. Note, such enterprise knowledge structures are bound to continuous change and refinement, e.g., due to changes to enterprise workflows. Semantic data structures, in contrast to relational data structures, can be extended at any time in SMW either by administrators or the users themselves without modifying previous contributions.

Retrieving information Users need to retrieve specific information from these enterprise knowledge structures. The machine-readable information stored with SMW gives more sophisticated possibilities of retrieving data from the wiki, other than traditional keyword searches. First, it can be specifically asked for certain properties of a page, e.g. the customer of a proposal. Second, all pages of a certain category having certain properties can be listed as an overview, including links to those pages, e.g, all products within a specific price range. Various result formats can be used, starting from simple tables to more advanced calendars, timelines, and maps. Also, *facetted search* is possible, for incrementally filtering lists of pages with keywords or property-ranges (see Figure 1 in a later section). Third, more complex but still user-friendly querying similar to the standard semantic web querying language SPARQL[4] is possible. The users can enter keywords, the system looks for connections between pages described with the keywords and lists those pages [HHMT09]. In SMW, these queries are possible through forms on special pages, but can also be embedded as *inline queries* in normal wiki pages.

Integrating external information Integrating external source allows to merge its content with existing enterprise knowledge structures, i.e. adding it as pages or properties,

[3] http://www.w3.org/TR/owl2-overview/
[4] http://www.w3.org/TR/rdf-sparql-query/

referencing it from other pages, and visualizing it in new ways. However, simply creating a page for each element within an external source and copying its data from there, will lead to difficulties searching and using this data. SMW provides a possibilty to tightly intergrate external sources. Enterprise knowledge seldomly is represented in RDF, but there are many tools[5] available to transform the formalized enterprise knowledge into RDF, which then can be mapped to specific knowledge structure elements in the wiki, e.g. categories, single pages, and properties. The more structured this external information is, the more this transformation can be done automatically. To deal with redundancy, it is possible to allow users to refer to, and comment upon external sources in the wiki, while changes may be undertaken only through the original systems and tools.

Not only data sources within an enterprise but also sources in the web such as Freebase or other Semantic MediaWiki installations can be integrated, this way. A growing number of web services offer data in RDF[6]. Using explicit bindings to such externally stored sources, wiki pages could be easily enriched with their data. The users may integrate new external sources by themselves, although this can be restricted, e.g, for security reasons.

Improving data quality One of the most useful features of a semantic wiki is its ability to perform consistency checks on the enterprise knowledge structures represented within the wiki and to indicate data quality problems. This provides a means to identify missing or incorrect information, which applies to both genuine wiki content and content from external sources. Users may not directly correct the latter, but they can rate it, and comment on it for revision. Besides the possibility that users detect inconsistencies, some checks can be performed automatically. Deduction methods on the enterprise knowledge structures can provide insights about the wrong usage of categories, pages, and properties [Vra09]. Most of such errors cannot be automatically repaired, but at least, made visible to the users or administrators. For example if the imported data contains information about a proposal with customer X and a wiki page exists about X, which is not a member of the customer category, adding that page to the category can be automatically suggested to the administrator.

Interplay with other enterprise tools Also, enterprise knowledge structures reach full potential if they are not stored in an isolated data silo but can be accessed from other enterprise tools. SMW allows not only to integrate external sources in a standard and machine-readable manner, but generally to use external tools to input from and output to the wiki. The content of a semantic wiki can be extracted as RDF, as well as many other structured data formats, e.g., JSON, vCard, and BibTeX. Results of queries can be regularly checked for new pages or for modified properties and published as RSS-feeds or send per e-mail. Using HTTP requests to the wiki, external tools such as Microsoft Word can access, add, or modify pages and properties.

Semantic technologies cannot generally fulfil the non-functional requirements of our scenario. Those are implementation-dependent and therefore will be described in the next

[5] `http://www.w3.org/RDF/`
[6] `http://www.linkeddata.org/`

section, where we evaluate our solution.

4 Implementation of a proposal development workspace

As it is our aim to have general guidelines for applying our solution, we do not depict differences between the wikis we realized for our three case studies, but focus on the lessons learned from the three implementations and describe and explain it as one. Our implementation is based on the open-source wiki software MediaWiki and its semantic counterpart Semantic MediaWiki, which have been augmented with a series of general-purpose extensions developed by the community[7] and custom extensions tailored to the needs of the proposal development scenario.

Users access the wiki from the intranet using a personalized and enterprise-wide login, realized through *LDAP Authentication* extension. Pages are cached until they are changed or reprocessing is explicitly requested. This way, SMW leads to not much loss in performance in comparison to a pure MediaWiki installation [HE10].

Creating proposal development information As a workspace for proposal development the wiki supports the entire life cycle of a proposal. For that, we have developed an ontology describing the proposal development structures. It contains categories of pages (classes) such as proposal, team, person, customer, customer issue, discussion question or answer, product, and event. Each of these categories are further defined through properties. For instance, a proposal has one or more proposal sections, a team of persons, a customer with customer issues, a monetary value, and offers one or more products; a customer issue can be discussed through questions and answers; a customer is related to an industry sector; and several meetings are held for a proposal. The ontology has been developed together with potential users, but also by reuse of available enterprise vocabularies and ontologies such as "The Enterprise Ontology" [UKMZ98]. We expect that continuous refinement of the ontology is necessary, also after an official launch of the system, and will mainly be done by administrators.

Each relevant element of a proposal is represented in the wiki as a dedicated page or part of a page. Users can easily create pages and property-value pairs through the *Semantic Forms* extension. Adding properties to proposal elements supports auto-completion, checkboxes, radio buttons, and other value selection widgets such as mini-calendars and map views. Furthermore, the extension *Header Tabs* facilitate the realization of forms that are similar in their appearance to the rendered pages. Discussions, e.g, about customer issues, have a structure similar to forum applications.

Retrieving proposal information In order to provide users with an overview of the proposal workspace, pages explaining and listing a particular aspect of proposal development

[7] openly available from `http://www.mediawiki.org/wiki/Extension_Matrix` (MediaWiki) and `http://semantic-mediawiki.org/wiki/Help:Extensions` (SMW)

are linked from the wiki main page. Administrators define queries and visualizations, e.g. listing all open proposals, all high-priority customer issues, and all products of a certain category. The users themselves can issue queries and store them on wiki pages, although we expect more use of keyword based searches that, with *AskTheWiki* [HHMT09] extension, still exploit the wiki's structure. Figure 1 shows the workspace customized through *Halo* extension, and an anonymized example of facetted search using *Exhibit*.

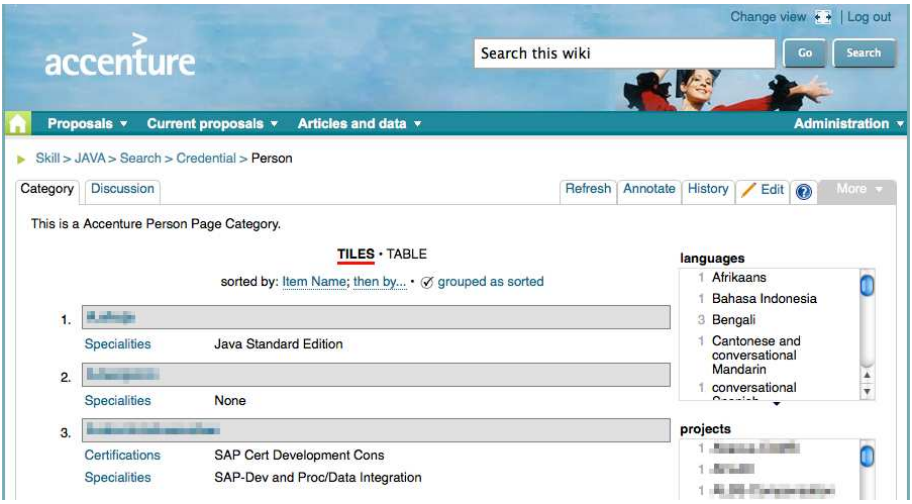

Figure 1: Anonymized wiki content displayed through facetted search.

Unnecessary searching and browsing is reduced by sharing relevant information between wiki pages. For instance, the property "hasCustomerVision" is not only shown on the particular customer page but also on proposal pages concerning this customer.

Integrating proposal-related information The proposal development structures established in the wiki were partly populated by content from external sources. We integrated externally created information about customers, industry sectors, offered products, time and pricing, people and their involvement in proposals, as well as status and result of proposals. This data mainly comes from relational databases or excel sheets, and is preprocessed, e.g., different date formats are matched.

Some of this information is first transformed in an RDF-compliant format, then imported as pages such as proposals, or properties such as a customer vision. Other imported wiki content such as person identification number is not supposed to be redundantly stored and open to changes. This information is only visualized, allowing users to refer to, and comment upon it in the wiki, while changes may be undertaken only through the original source. Although promising for the future, so far, we have not found useful Linked Open Data for proposal development.

Improving proposal data quality The system does not only allow input and retrieval of new or existing information, it also ensures long-term data quality. Integrating various external data sources into the wiki makes implicit connections between them explicit. For instance, proposal descriptions may contain team member names, that provide direct links to pages of those persons. Misspelled names can be easily identified through links to non-existing pages. Correct abbreviations or accronyms can be included as synonyms. More data quality problems are identified and corrected by the administrators when they export the data as RDF using *RDFIO* extension and compare its consistency with the proposal development ontology. That way, wrong properties can be detected and made visible to the users, e.g., an event having a monetary value; a customer being located in a proposal; and a due date containing a number.

Interplay with tools for proposal development For automatic notifications about certain changes in the wiki, users can subscribe to RSS-feeds. Those publish information about new pages, e.g., such as new customer issues with a high-priority, or more fine-grained information about modified properties such as the monetary value of a proposal. Such information can also be displayed as a widget on other company intranet sites, and, using *Semantic Notifications* extension sent per e-mail.

When finally creating the actual proposal document within Microsoft Word, the users can access the structured information in the wiki through the *WikiTags* extension. If a particular proposal element is mentioned in the document, Word automatically underlines it and provides a wiki-based context menu for it, e.g., for fast copy-paste-like insertion of properties stored in the wiki (Figure 2). Vice versa, for new expressions not yet used in the wiki, it is possible to create wiki pages, properties and links directly from the office application. Once additional information about these is collaboratively assembled in the wiki, the results can be used in the Word document.

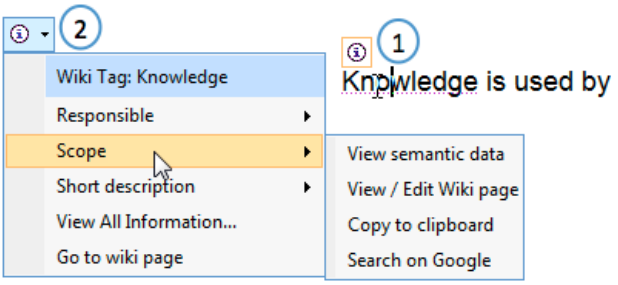

Figure 2: Microsoft Word underlining a word (1) and providing a wiki-based context menu (2).

Within the ACTIVE project, the implemented solution was presented to the potential users, who have confirmed that it fulfils the requirements. Recently, trials have been initiated for evaluating the solution with actual users.

5 Related work

Not much public information can be found on technical challenges of enterprise wikis. Danis and Singer [DS08] present a detailed study of the differences in wiki utilization in corporate, educational and public settings, but without going into detail on technical aspects. If any, only general solutions to technical challenges can be found. For instance, deciding what hardware or service to use, and offering features such as a more simple file upload and an indexing of attachments for built-in search [Arc10].

The social software market offers various enterprise systems [DRBM09] that among other things provide wiki solutions. IBM Lotus Connections and Microsoft's SharePoint Server, as market leaders, still lack strong wiki functionality. Atlassian Confluence and TWiki may provide this, however, they do not give details about their approaches to technical challenges. KiWi is a wiki solution that combines Web 2.0 and Semantic Web technologies similarly to SMW and can be used to build enterprise applications [SMSK10].

6 Conclusions

In this work, we have proposed and implemented a semantic-wiki-based solution for introducing a wiki in a typical enterprise scenario. It fulfils the technical requirements and beyond that opens up new opportunities for the enterprise, e.g., exploitation of the growing Linked Open Data and seamless interplay between the wiki and external tools.

We now conclude with some final remarks and possible future work. Due to the large number of extensions, choosing the most appropriate ones is a tedious task, that could be improved by further empirical studies on the utility of existing or possible extensions in specific application scenarios. For now, we have evaluated the hypothesis that the technical challenges can be resolved by a semantic wiki. We have not evaluated whether the requirements acually were representative and whether our solution will bring the expected benefits. Measuring its impact is under work in recently initiated trials. As an example, the time the user spend with the wiki, either directly or through external tools, could be an indicator for sucess. Still, evaluating the success of a wiki is not an easy task, due to many confounding success factors. We have not tried to fulfil the requirements with other technologies than Semantic MediaWiki and, thus, cannot say much about comparison to other systems, be they semantic wikis or not. As intended, we solely argue that semantic wiki technologies clearly expand the capabilities of knowledge workers for managing enterprise knowledge structures.

Acknowledgements

Work presented in this paper has been funded by the EU IST FP7 project ACTIVE under the grant 215040.

References

[AKTV07] Anupriya Ankolekar, Markus Krötzsch, Thanh Tran, and Denny Vrandečić. The two cultures: mashing up web 2.0 and the semantic web. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 825–834, New York, NY, USA, 2007. ACM Press.

[Arc10] Bill Arconati. Factors for the success for Wikis 2: Organization is the key (blog post), 2010. `http://blogs.atlassian.com/confluence/2010/06/factors-for-success-for-wikis-2-organization-is-the-key.html`, Last visit 07/11/2010.

[DRBM09] Nikos Drakos, Carol Rozwell, Anthony Bradley, and Jeffrey Mann. Magic Quadrant for Social Software in the Workplace. Gartner RAS Core Research Note G00171792, Gartner, 10 2009.

[DS08] Catalina Danis and David Singer. A wiki instance in the enterprise: opportunities, concerns and reality. In *CSCW '08: Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 495–504, New York, NY, USA, 2008. ACM.

[HDW10] Lester J. Holtzblatt, Laurie E. Damianos, and Daniel Weiss. Factors impeding Wiki use in the enterprise: a case study. In *CHI EA '10: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 4661–4676, New York, NY, USA, 2010. ACM.

[HE10] Daniel Herzig and Basil Ell. Semantic MediaWiki in Operation: Experiences with Building a Semantic Portal. In *9th International Semantic Web Conference (ISWC2010)*, November 2010.

[HHMT09] Peter Haase, Daniel M. Herzig, Mark Musen, and Duc Thanh Tran. Semantic Wiki Search. In *6th Annual European Semantic Web Conference, ESWC2009, Heraklion, Crete, Greece*, volume 5554 of *LNCS*, pages 445–460. Springer Verlag, Juni 2009.

[KVV+07] Markus Krötzsch, Denny Vrandečić, Max Völkel, Heiko Haller, and Rudi Studer. Semantic Wikipedia. *Journal of Web Semantics*, 5:251–261, September 2007.

[SMSK10] Rolf Sint, Mark Markus, Sebastian Schaffert, and Thomas Kurz. Ideator - a collaborative enterprise idea management tool powered by KiWi. In *5th Workshop on Semantic Wikis: Linking Data and People (SemWiki2010)*, 2010.

[STW+ar] Elena Simperl, Ian Thurlow, Paul Warren, Frank Dengler, John Davies, Dunja Mladenic, Jose Manuel Gomez-Perez, Carlos Ruiz Moreno, and Marko Grobelnik. Overcoming Information Overload in the Enterprise: the ACTIVE Approach. *IEEE Internet Computing*, November-December 2010 (to appear).

[UKMZ98] Mike Uschold, Martin King, Stuart Moralee, and Yannis Zorgios. The Enterprise Ontology. *Knowl. Eng. Rev.*, 13:31–89, March 1998.

[VK06] Denny Vrandečić and Markus Krötzsch. Reusing Ontological Background Knowledge in Semantic Wikis. In Max Völkel and Sebastian Schaffert, editors, *Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics*, Workshop on Semantic Wikis. AIFB, ESWC2006, June 2006.

[Vra09] Denny Vrandečić. Towards Automatic Content Quality Checks in Semantic Wikis. In *Social Semantic Web: Where Web 2.0 Meets Web 3.0*, AAAI Spring Symposium 2009, Stanford, CA, March 2009. Springer.