TOOLS TO FOSTER SEMANTIC-BASED COLLABORATION
A Knowledge Management Approach Based on a Semantic Wiki and Personal Ontologies

Flavio De Paoli and Marco Loregian
Dipartimento di Informatica, Sistemistica e Comunicazione, University of Milano-Bicocca
via Bicocca degli Arcimboldi 8, 20126 Milano, Italy

Keywords: Collaboration, Knowledge Management, Semantic Wiki, Ontology Editing.

Abstract: Increased mobility and team distribution may hinder collaboration in organization by hampering spontaneous communication in workplaces. Tools are required to be integrated and not intrusive with respect to users' activities, in order to provide effective support to workers for the development of joint projects. In this paper we present the enhancement of a knowledge-based collaborative system by means of a Semantic Wiki, which is a wiki exploiting ontologies to create machine-processable annotation of documents. Moreover, we present a simple tool to edit personal ontologies, that can be shared with other team or organization members.

1 INTRODUCTION

Being composed of several, often parallel activities and tasks, business processes rely on active collaboration among participants. Several approaches have been proposed for business process management (BPM), mainly focusing on issues related to coordination among entities (van der Aalst et al., 2003). However, besides the planning of processes, *spontaneous* communication among peers has an important role: it usually takes place when people meet each other in physical settings, casually or on purpose. With increased mobility of users and distribution of teams, such a possibility is at risk, and technological means seem to be required to overcome the possible inability of encountering colleagues and to foster spontaneous participation in the evolution of projects.

The aim of this work is to propose the integration of techniques that have already been proved successful in world-scale situations into the smaller setting of modern enterprises. In particular, we aim at presenting how to integrate Semantic Web and wiki technology into a Knowledge Management (KM) system.

A KM system is a collaborative environment supporting the workers of an organization in exchanging knowledge by providing an environment to let them share documents and supplying techniques to associate a meaning with them — usually adopting ontologies and profiling mechanisms. The reference platform of this paper was designed as the result of studies conducted on real settings to analyze the role of (informal) communication and the way it is used in articulating work tasks (Agostini et al., 2005).

In this paper, we present a set of features promoting the involvement of people in creating and exchanging knowledge by providing them with an easy way of publishing group knowledge. The chosen approach is to combine the definition and sharing of personal ontologies with the creation of profiles and annotations of documents using a wiki interface (Leuf and Cunningham, 2001). A personal ontology is composed of a set of terms and relations that are dynamically selected from available ontologies, possibly augmented with user-defined semantic extensions. A tool to facilitate the definition and the evolution of personal ontologies, their use in common activities, such as Web browsing, and KM specific, such as document profiling, is described.

2 FEATURES TO ENHANCE COLLABORATIVE KM

This section starts with the outline of the general characteristics of the reference KM platform used in our

De Paoli F. and Loregian M. (2007).
TOOLS TO FOSTER SEMANTIC-BASED COLLABORATION - A Knowledge Management Approach Based on a Semantic Wiki and Personal Ontologies.
In *Proceedings of the Third International Conference on Web Information Systems and Technologies - Web Interfaces and Applications*, pages 302-307
DOI: 10.5220/0001286603020307

experiments; next, the Semantic Wiki designed and developed to support collaboration is presented with its features and architecture; finally, a tool to edit the personal ontologies to be used in the KM system is described.

2.1 The Reference KM System

MILK[1] is a Knowledge Management system that allows a community of persons to store, organize and retrieve documents in a multimodal and multichannel way. The system relies on a server-based architecture encompassing a KM engine (the Metadata Management System, MMS) that can be connected to heterogeneous archiving systems.

At KM level, MILK defines a profiling mechanism that unifies knowledge descriptions associated with documents, people, projects and communities (collectively named *elements*). Any kind of element can be included in the same indexing and retrieval process. So, for example, given the profile of a project, it is possible to collect information about any other project that has similarities with the current one, know who is interested in the same subject, and identify documents regarding the same topics. The profiling process is driven by ontologies that describe specific domains: ontologies can be used to edit elements' metadata, and are then exploited by the MMS to compute correlations between elements — i.e., to identify related topics (Boselli et al., 2003) .

At interaction level, the aim of the system is to provide users with various contextualized interfaces and interaction mechanisms that fit different devices and working situations, supply users with different views over the same contents to promote awareness and learning according to the activity a user is undertaking (Agostini et al., 2005; Agostini et al., 2003). In other words, user interfaces are requested to provide multimodal interactions according to different user terminals.

The need of supplying multimodal and multichannel presentation of documents — that is, to be able to present the same document in different formats according to the most suitable device currently available to the user — led to a key choice in the information architecture of the system. Documents are composed of a collection of files, each one proposing a *representation* of the actual content of a document. For example, for each document might, or rather should, exist a representation that is well-suited for a regular PC (e.g., the full *pdf* version of a paper), one for

[1]Multimedia Interaction for Learning and Knowing, IST-2001-33165, `http://www.cootech.disco.unimib.it/milk/`

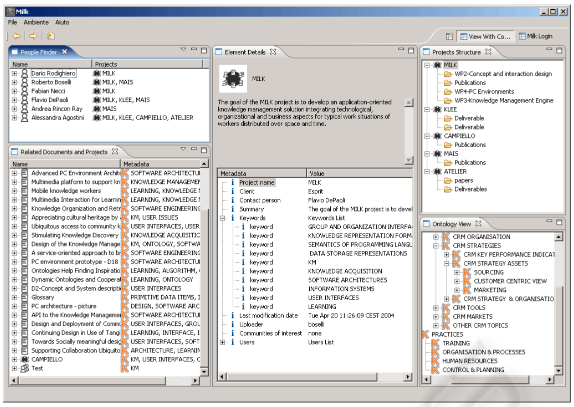

Figure 1: The MILK desktop client GUI, showing the "View With Context" presentation.

large displays (e.g., a slide-show presenting the content), and one for handheld devices (e.g., the textual abstract of the paper). It is important to remark that all these files should be considered as being the same document or, at most, different views of the same document. Thanks to profiling, the system is able to manage the different representations as "attachments" to the same element.

The most general representation of a document is its profile, a set of metadata that is stored by the system, that can be edited by users using a (standalone or web-based) client, and that can be enriched by means of Semantic Web technology (RDF) and successful interaction techniques (wiki-style collaboration). Ontologies fitting users' needs are essential to achieve good results in supporting work practices.

In order to easily exploit ontologies in element profiles, the GUI provides a standard way to visualize and use ontology terms, as keywords, in document profiles: a panel displays such terms in a classic tree fashion from which a term can be added to an element profile by simply dragging and dropping it. Otherwise, terms can be manually entered by typing them in the appropriate form fields. Typed terms can be part of the ontology or new keywords. In the latter case, they are collected and proposed to the authorized administrators of the system as candidate terms for a reference ontology that is shared organization-wide.

2.2 Collaborative Knowledge Creation

The system helps people to manage information and easily access the right sources for creating new contents (Agostini et al., 2003). In most KM systems, users are enabled to *search* repositories to get information; instead, MILK supports users to *discover* in-

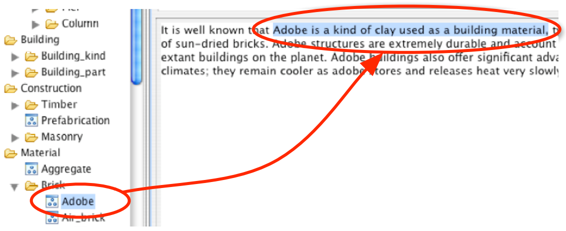

Figure 2: A detail of the Wiki perspective embedded in the KM client.

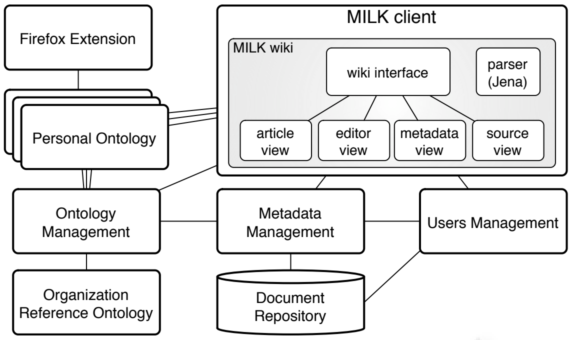

Figure 3: The architecture of the system: only elements relevant to the wiki perspective are shown.

formation while they are using the system. The discovery is supported by automatically computing and presenting information that are related to what the user is doing and displaying on the screen. Figure 1 illustrates the concept in a prototype (desktop PC) interface that displays a document surrounded by elements that are related to the content of that document. This interface style is named "view with context" to say that any displayed element is surrounded by its context.

The current GUI is a standalone application, built on Eclipse Rich Client Platform[2]. The advantage of this choice is to deliver a (cross-platform) tool that is easily extensible, via plug-ins, to allow for rapid proofing of new ideas.

The interface is structured in *perspectives* that collect related *views*. For example, the "View with Context perspective" is composed of a main view surrounded by smaller panels displaying ontologies, related elements, project structure, and available people (Fig 1). The main view (i.e., the so-called *edit metadata* view) is essentially form-based, meaning that metadata is entered in profiles by filling in the fields of a form. In Computer Supported Cooperative Work (CSCW) literature, this task has been widely recognized to be tedious. In particular, this kind of interaction is not perceived as a collaborative activity. Forms suggest a rigid conception of profiles, whereas the Metadata Management System was designed for more flexibility (e.g., to deal with heterogeneous elements).

On the other hand, the popularity of tools such as Wikipedia (Giles, 2005) shows that wiki-style interaction could be the key factor to allow users to generate new contents by completing existing element profiles with information that can be easily edited (with elementary wiki syntax) and exploited, for example, to build richer views on the knowledge base.

For these reasons, we are testing a "Semantic Wiki perspective" in the RCP client (Figure 2). Users can annotate elements' profiles as free-text, and we claim it to be a *semantic* wiki because of the enhancement it carries in knowledge processing, being tightly integrated with system ontologies. The definition of a wiki representation enhances the content management capabilities of the system, for example to support more precise computation of the view with context, taking annotations into account.

At a basic level, annotations can be associated with a document by simply selecting words or sentences in the text area and then dragging one or more terms from the ontology view on them. By this action, a RDFa[3] annotation is created in the document, and the user perceives it through the change of the annotated portion of the text that turns underlined.

The semantic wiki was designed according to a simple, yet consolidated architecture (Oren et al., 2006) (Figure 3). In the user interface, four main views are currently available:

- the *article* view is a read-only display of the wiki representation of the document. With this view users can get general view over the annotations on a document: annotations are associated with underlined text. Using mouse roll-over tooltips it is possible to view annotation details. Using an HTML template, with XSLT and CSS technologies, the wiki description can be rendered on a range of devices or exported in various formats, with various appearances;
- the *editor* view presents a compact editable view of the document: the text and the annotation can be modified by picking terms from the available ontologies. Also in this view, annotations are only associated with underlined text;
- the *edit metadata* view collects the list of generic metadata and the annotations in a form: metadata

[2] http://www.eclipse.org/rcp/

[3] http://www.w3.org/TR/xhtml-rdfa-primer/

Figure 4: The Mozilla Firefox extension, and the main interaction steps to define a personal ontology by adding a new term.

can thus be quickly edited and the changes are automatically inserted in the wiki file;

- the *source* view allows the user to directly edit the source code of the wiki file (i.e., RDFa annotations): tags are shown and can be directly edited in the body of the document.

When a document is saved, the MMS is in charge of parsing the wiki file (the wiki perspective relies on Jena APIs to generate documents) to extract information and update the document profile accordingly. The RDFa representaion is stored in the repository along with other representations and documents.

2.3 How to Create Personal Ontologies

In its original release, the MILK system was limited to rely *only* on a centralized ontology. Such ontology was designed in different versions to accommodate the needs of the test settings in which the system was deployed. Generic APIs and graphical user interfaces were implemented to directly manipulate ontologies. Ontology maintenance required administrators that were familiar with basic principles of knowledge engineering and with MILK technology.

A goal of our work is to design features that allow individuals to easily define personal ontologies, possibly grounded on those published on the Web, that can be shared with colleagues, instead of only allowing users to type new keywords. Our approach is to integrate an ontology creation process with the most common desktop activity: *Web browsing*.

In the everyday activity, we use the Web to extend our knowledge by learning from documents and from terms exploited in the searching. We are conducting experiments that aim to extend the system beyond the basic activities on shared ontologies and repositories to include the terms used during browsing as personal ontology terms. We are testing an extension of Mozilla Firefox that monitors the queries sent to search engines, e.g., Google (Figure 4). The extension is called OntoMarks and can be downloaded from the following URL: `http://www.cootech.disco.unimib.it/ontomarks/`.

The user interface of the browser extension consists of a sidebar with an upper "permanent" panel, and a lower "temporary" panel. Whenever a Google page is loaded, the keywords used in the search are extracted by parsing the page URL and added to the bottom panel of the sidebar (step 1 in Figure 4). By right-clicking on any of the keywords currently displayed, the user can search the term again with Google (like with a bookmark) or make the keyword part of a permanent set (step 2 in Figure 4). Searches with multiple keywords are split, but expressions between quotes are kept as a single keyword. The list of keywords in the panel is temporary, and it is reset every time Firefox is started.

When a keyword is selected to be included in the personal ontology, a query can be sent to Swoogle[4], via its Web Service APIs, to find out if the term is defined by available ontologies. A new browser tab is opened to show the matching results and let the user select the most appropriate term by reading available descriptions. The insertion in the personal ontology can be postponed, and the found terms are temporarily saved and accessible via a contextual menu that can be open by right-clicking on the already searched term. A colored dot shows the status of a term: red for terms that have not been searched or that have not been found on Swoogle, green otherwise.

If a term from an existing ontology is chosen, then a complete copy of its definition (description, properties, relations) is added to the personal ontology, so to be available for future use, for example when the ontology is shared with other team/organization members (Section 2.4)

Besides using terms as bookmarks for Google searches, the general utility of creating a personal ontology is to use it in KM systems such as MILK (i.e., by loading and using it in the GUI) to better organize personal knowledge or retrieve information from a shared repository: needless to say, it is the vision of the Semantic Web (Berners-Lee et al., 2001).

For this reason, this plug-in is compliant with Semantic Web technologies: the upper panel relies on a RDF file that permanently stores a user's personal ontology, that can be loaded and managed by the client GUI as well as by any other RDF-based application.

[4] `http://swoogle.umbc.edu/`

2.4 How to Share Personal Ontologies

MILK's KM engine can supply contextualized access to document repositories, through different interfaces, and can load personal ontologies from local hard drives and from the Internet (via URIs). By doing so, users can request the MMS for customized views on the knowledge base. However, to support cooperation, personal ontologies might be needed as shared resource (in parts or as a whole) since they are referenced by profiles and wiki annotations. Moreover, they can be a vehicle to spread new knowledge.

The system supports access to a shared reference ontology, that is maintained by administrators and to which all organization users can contribute by adding new terms (De Paoli and Loregian, 2006b). More precisely, users can propose new candidate terms, by simply using them in document profiles, and administrators have the task of accepting them. Administrators have actually the more complex role of keeping the reference ontology consistent with the organization activities, and to supply additional information (e.g., relations among terms) in order to allow the MMS to effectively compute correlations among terms and, therefore, to enable mechanisms such as the already mentioned *view with context*.

Annotations can carry different amounts of information, according to the technique adopted by the users: (i) if a new term is simply typed in a proper form field, no additional semantic information (i.e., definitions, properties, and relations with other terms) is provided. In this case, the keyword is a simple candidate term, and administrators have to possibly provide its semantics; (ii) if a term is dragged and dropped from a personal ontology, an ontology fragment, rather than just a term, is proposed as a candidate. By fragment we mean a term with all its known relations, possibly along with other involved terms. In this case administrators can take advantage of existing references, or they can use the original ontology (from which the term was chosen and taken by the user) as a complete reference.

Annotations can help administrators' tasks by providing the context in which a term was used, and also end-users can extend their knowledge of a specific topic by looking for annotations referring to a specific term.

3 RELATED WORKS

The work presented in this paper took inspiration from other efforts made in developing semantic wikis. Among others, SEMPERWIKI (Oren et al., 2006), and IkeWiki (Schaffert, 2006). While other wikis have been designed as standalone tools, with the purpose of creating homogeneous, yet processable knowledge bases, the project described here aims at going a step further. With respect to KM, it seems rather unrealistic to consider wikis apart from other existing and already adopted collaborative tools, or as a new technology to be applied only when starting to build new collaborative knowledge bases. Thanks to the simplicity with which annotations can be added to existing documents, wikis should instead be integrated with technologies in use. In this paper we have discussed the integration with the MILK system, which relies on a file system managed by an FTP server. An advantage of this technique is that the integration can be fully transparent to end users, which may decide to ignore the feature. Other integration experiments could be carried on with respect to other existing document management platforms, such as BSCW (formerly adopted in MILK) (Agostini et al., 2005).

Moreover, thanks to the reuse of semantic terminologies — enabled by the ability of the MILK client to load ontologies from various locations and by the Firefox extension to edit personal ontologies — users can be supported all across their main working and learning activities without forcing them to drop the tools to which they are familiar with.

The browser-extension approach is already widely in use to enrich the browsing experience with additional information. For example, an extension called Semantic Turkey has also been developed to edit ontologies within the browser (Griesi et al., 2006). Differences with our approach arises in the aims and in the architecture, as the extension is essentially the GUI of the whole knowledge management system. While our approach exploits a lightweight, client-side implementation of the bare ontology editing functionalities, the Semantic Turkey system relies on server-side processing and storing and unifies the process of editing the ontology (i.e., defining terms and relations) with the creation of annotations in Web pages, which are stored as instances of the same ontology and manipulated adopting a Sesame database. In our approach, the latter part is managed by the MMS independently from the extension, and the semantic wiki is used to create annotations. More important, the Semantic Turkey essentially allows for the collaborative editing of a reference ontology, which differs from our aim of allowing for the definition of individual ontologies that can be used collaboratively, when needed, in a proper way.

From a cognitive and interactive perspective, the process of creating and editing a personal ontology with the extension presented in Section 2.3 is similar to the organization of regular Web bookmarks (Abrams et al., 1998). Such activity is often considered unnecessary as long as the number of bookmarks is limited, and all the favorite URLs can be kept together unsorted, but most Web users tends to develop personal bookmark hierarchies (with folders and subfolders by topic) to sort their references. The result of such operation, which usually takes only a few minutes, is an explicit representation of the interest of a person. Additionally, our browser extension can be used to express such conceptualization in a standard language, thus, ready to be shared with other users. It is not just *social bookmarking* or annotation (Millen et al., 2006; Wu et al., 2006), in that the generated ontology remains personal until parts of it are willingly used in a collaborative setting and, thus, shared.

4 CONCLUDING REMARKS

In this paper we have presented the ongoing work to design a KM tool integrated with the Semantic Web philosophy: other experiments have already been made in this direction (De Paoli and Loregian, 2006a). We have outlined how a semantic wiki can be introduced and exploited in a knowledge management system to effectively support explicit knowledge creation. The advantages of adopting such an approach is to promote the publishing activity by letting users write textual descriptions augmented with semantic annotations, and to promote collaborative profiling.

We believe that adding personal annotations to existing documents, by using personal ontologies, should help involving users in creating new organizational knowledge. Moreover, the possibility of discovering new terms and relations among documents is a rewarding additional feature.

This paper presents the integration of sample features in an open platform — our reference KM system — and we are currently evaluating users' behavior and responses. Our future work consist in the design and development of new features that integrate popular tools to enhance shared and cooperative features. For example, we are currently testing the integration of RSS feeds, so that users can edit publish links or documents having their favorite news sources as a reference.

REFERENCES

Abrams, D., Baecker, R., and Chignell, M. (1998). Information archiving with bookmarks: personal web space construction and organization. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 41–48, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co.

Agostini, A., Albolino, S., Boselli, R., De Michelis, G., De Paoli, F., and Dondi, R. (2003). Stimulating knowledge discovery and sharing. In *GROUP'03*, pages 248–257. ACM Press.

Agostini, A., Albolino, S., De Paoli, F., Grasso, A., and Hinrichs, E. (2005). Supporting communities by providing multiple views. In *Proceedings of the Second Communities and Technologies Conference 2005*, pages 437–456. Springer.

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American*, 284(5):34–43.

Boselli, R., Dondi, R., and De Paoli, F. (2003). Knowledge organization and retrieval in the MILK system. In *SEKE 2003*, pages 372–376.

De Paoli, F. and Loregian, M. (2006a). Enhancing Semantic Interoperability in Collaborative Systems. In *Proceedings of the Eighteenth International Conference on Software Engineering and Knowledge Engineering (SEKE 2006)*, pages 29–34. Knowledge System Institute Graduate School.

De Paoli, F. and Loregian, M. (2006b). On Integrating a Semantic Wiki in a Knowledge Management System. In Tumarello, G., Bouquet, P., and Signore, O., editors, *Proceedings of SWAP 2006, the 3rd Italian Semantic Web Workshop, Pisa, Italy, December 18-20, 2006*, volume 201. CEUR-WS.org.

Giles, J. (2005). Internet encyclopedias go head to head. *Nature*, 438:900–901.

Griesi, D., Pazienza, M. T., and Stellato, A. (2006). Gobbleing Over the Web With Semantic Turkey. In Tumarello, G., Bouquet, P., and Signore, O., editors, *Proceedings of SWAP 2006, the 3rd Italian Semantic Web Workshop, Pisa, Italy, December 18-20, 2006*, volume 201. CEUR-WS.org.

Leuf, B. and Cunningham, W. (2001). *The Wiki Way: Collaboration and Sharing on the Internet*. Addison-Wesley Professional.

Millen, D. R., Feinberg, J., and Kerr, B. (2006). Dogear: Social bookmarking in the enterprise. In *CHI '06*, pages 111–120, New York, NY, USA. ACM Press.

Oren, E., Breslin, J. G., and Decker, S. (2006). How semantics make better wikis. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 1071–1072, New York, NY, USA. ACM Press.

Schaffert, S. (2006). IkeWiki: A Semantic Wiki for Collaborative Knowledge Management. In *1st International Workshop on Semantic Technologies in Collaborative Applications STICA 06, Manchester, UK*.

van der Aalst, W. M. P., ter Hofstede, A. H. M., and Weske, M. (2003). Business process management: A survey. In *Business Process Management*, pages 1–12.

Wu, X., Zhang, L., and Yu, Y. (2006). Exploring social annotations for the semantic web. In *WWW '06*, pages 417–426, New York, NY, USA. ACM Press.