Towards Social Webtops Using Semantic Wiki

Jie Bao, Li Ding, Deborah McGuinness, James Hendler
Department of Computer Science
Rensselaer Polytechnic Institute
Troy, NY, USA 12180
{baojie,dingl,dlm,hendler}@cs.rpi.edu

ABSTRACT

The fast-growing Web 2.0 applications promote explicit social structure and explosive data on the Web, and enable the realization of social webtops, where users use web applications to collaboratively organize and share online data. In this work, we show that semantic wiki technologies are suitable for building a social webtop. We also identify two key issues in realizing social webtop applications on semantic wiki and present our initial investigation with live demos: (i) popular concept modeling mechanisms used by webtop applications can be supported by semantic wiki; and (ii) provenance-aware data personalization mechanisms can be added as semantic wiki extensions to better support collaborative data management on a webtop.

Keywords

Semantic Wiki, Webtop, Concept Modeling

1. INTRODUCTION

The rapid growth of personal and community data on the Web demands effective *social webtops*, i.e., Web-based infrastructure for users to organize and share online data. Being the Web-based counterpart to a desktop system on a personal computer (PC) , a social webtop uses the Web for data storage and provides specific applications for seamlessly moving our daily data process applications onto the Web, e.g., Wikipedia reduces the need for encyclopedia software on PC, and Google Document offers a Web-based alternative to conventional word processing software on PC.

A successful social webtop requires many critical capabilities, such as structured data representation, smart data integration and propagation, provenance-aware social data access control, Web data persistency and preservation, and friendly data access UI. In this work, we focus on the following two critical data management requirements:

• The ability to support effective *data organization* on the Web. A social webtop should provide versatile concept modeling supports to facilitate web users creating, propagating, accessing, and consuming online data conforming to various application conventions.

• The ability to support *data sharing* on the Web. A social webtop should be aware of social provenance and support both *socialization* that promotes collaborative data generation and consumption, and *personalization* that ensures necessary data privacy protection and customization.

Among the approaches to social webtops, semantic wiki, according to our recent study, is a promising platform for building social webtops with low development and adoption cost. Our experiments also identified two critical data management gaps when building social webtop applications on semantic wiki: (i) the concept modeling choices of many social webtop applications, such as a blog and a social recommender, are not natively supported by semantic wiki's RDF modeling; and (ii) personalized data sharing supports, such as privacy protection and provenance tracking, are poorly supported by the current semantic wiki (e.g., there is no way to prevent a semantic wiki editor from retrieving sensitive data encoded in the triples stored by semantic wiki).

2. SOCIAL WEBTOP ON SEMANTIC WIKI

The recent development of semantic wikis greatly reduces the cost of developing and adopting social webtop applications. Inheriting the native collaborative data sharing support from conventional wikis, semantic wikis, such as Semantic MediaWiki (SMW, [1]), provide simple extensions to the wiki scripts to generate and query semantic annotations in wiki articles. Semantic wiki, according to our study, can not only be a general purposed tool for content management, but only a powerful workbench for building light-weight social webtop applications. The latter vision is illustrated in Fig. 1 where social webtop applications can be developed on internal and external online data using wiki-based concept modeling and programming power. Each layer gives a higher level of abstraction of the information on the layer below.

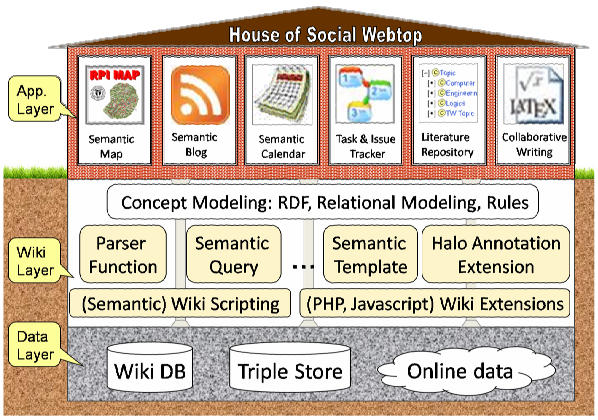

Figure 1: Semantic Wiki as an Application Infrastructure

• *The data layer*: it refers to internally stored data including the normal wiki contents and the RDF triples parsed from wiki articles, and imported online data such as Google Map and RSS feeds.

• *The wiki layer*: wiki provides collaborative and extensible programming facilities including the PHP-based wiki extension architecture and wiki scripting languages (such as semantic query and parser function) for a developer to control the generation, representation and propagation of wiki data. It hides many details of the data layer, e.g., whether data is stored in a relational database, an internal triple store, or an external triple store. Thus, a developer can focus on the conceptual model of data. As wiki scripts are also stored as wiki pages, this architecture has the inherent support for collaborative development and version control.

• *The application layer*: it provides social webtop applications, developed using wiki scripts, for normal web users to generate or consume information as if they are using Web 2.0 or desktop applications. This layer offers friendly form-based or graphical user interface for publishing, accessing and sharing data, and it hides wiki scripts from end users.

Our study also results in a handful semantic wikis using SWM including (i) *TW Wiki*[1] which servers as a semantic group portal, (ii) *RPI Map*[2] which mashes up with external Web data, and (iii) *TW Wine Wiki*[3] which maintains a collaborative knowledge base for an intelligent wine recommendation agent.

In the following two sections, we show our approaches to the two gaps between semantic wiki and social webtop.

3. DATA ORGANIZATION

Besides RDF modeling natively enabled by SMW, other concept modeling approaches are also widely used as design patterns in Web data organization. In what follows, we use TW Wiki to show the following three popular concept modeling approaches.

• *RDF Modeling.* SMW lets users use RDF to (i) model wiki pages as named RDF resources, such as *instances* (normal wiki pages), *classes* for grouping wiki pages (pages in the namespace Category), *properties* for describing and associating pages (pages in the namespace Property); (ii) annotate wiki pages with RDF triples; for example, on a person's page with the name 'John', triples such as [`John rdf:type Person`] and [`John foaf:knows Jane`] can be asserted.

• *Relational Modeling.* With the help of semantic template[4], wiki data can also be modeled like relational tables, where a template definition can be viewed as the schema of a relation table and each wiki page that consists an instance of the template can be viewed as one row of the relational table. For example, TW Wiki intensively uses relational modeling to track information about project development issues including (i) the description of an issue, (ii) the persons who open, close or are assigned to the issue, (iii) the projects associated with the issue, and etc. Here, we modeled a few relations, i.e., `Issue`, `Person` and `Project` using semantic templates, and use template attributes of `Issue` to capture the relations, e.g., `openedBy`, from `Issue` to `Person` and `Project`.

[1]http://tw.rpi.edu

[2]by Jin Guang Zheng. http://map.rpi.edu

[3]by James Michaelis. http://onto.rpi.edu/wiki/wine

[4]http://www.mediawiki.org/wiki/Extension:Semantic_Forms

• *Rule Modeling.* Using parser functions[5] and semantic query, semantic wiki also supports simple rule modeling. For example, TW Wiki uses rule modeling to infer indirect affiliations of a person: a wiki script based rule may use the facts [`Bob memberOf TWGroup`] and [`TWGroup partof RPI`] to derive the fact that [`Bob memberOf RPI`].

4. DATA SHARING

Personalization issues arise frequently in our wiki development experience. The Wikipedia model that minimally restricts read and write privilege may not fit all social webtop applications. For example, TW Wiki may need to hide internal group meeting notes from non-group members while keeping all publication publicly visible. In what follows, we discuss three important personalization issues with some initial solutions:

• *Ontology Personalization.* A nice feature offered by semantic wiki is ontology based knowledge organization, where wiki contents are organized by the category taxonomy. According to well-known difficulties in ontology convergence, open nature of the Web, and our real experiences (even two experts cannot fully achieve full agreement on a common ontology), one common ontology can hardly accommodate the needs from all users. Therefore, TW Wiki offers semantic annotation approach that maintains a shared ontology and lets users selectively use or exclude concepts from that ontology to create their personalized ontologies. Currently, we are investigating a wiki extension that lets users query and present data using personalized ontologies.

• *Privacy Protection.* In order to maintain both public and private data on social webtops, privacy protection should be supported by semantic wiki. We have experimented several simple approaches on TW Wiki, e.g., the access control policy on the semantic blog component: every user can write a feedback but not a post on other users' blogs. We are currently working on a policy language and a privacy-preserving query engine for protecting facts and semantically inferred facts.

• *Provenance.* Provenance information maintained by wiki tracks who updated which wiki page at when. By associating this provenance information with the semantic data, such as category information, it is easy to collect and analyze wiki users' activities and interests. TW Wiki has implemented an initial semantic history wiki extension that converts the wiki revision history into RDF data. Currently, we are improving this wiki extension to support provenance query and support provenance aware privacy protection.

5. CONCLUSIONS

Our initial investigation on semantic wiki has shown that it is a suitable and low cost platform for building social webtop applications. We also identify two critical issues learned from our experiments and suggest solutions in certain use cases. Our future work will further improve our solutions to fill the data organization and data sharing gaps between semantic wiki and social webtops.

6. REFERENCES

[1] M. Krotzsch, D. Vrandecic, and M. Volkel. Semantic mediawiki. In *International Semantic Web Conference*, pages 935–942, 2006.

[5]http://meta.wikimedia.org/wiki/ParserFunctions