OntoWebber: A Novel Approach for Managing Data on the Web

Yuhui Jin, Sichun Xu, Stefan Decker, Gio Wiederhold
InfoLab, Stanford University, Stanford, CA 94305, USA
E-mail: {yhjin, xusch, stefan, gio}@db.stanford.edu

Abstract

OntoWebber is a system for managing data on the Web with formally encoded semantics. It aims at solving the problems current technologies are confronted with, namely, the reusability of software components, flexibility in personalization, and ease of maintenance for data intensive Web sites. Based on a domain ontology and a site modeling ontology, site views on the underlying data can be constructed as site models. Instantiation of these models will create the browsable Web site. And the manipulation of the site models helps to reduce the high effort for personalizing and maintaining the Web site. In this paper we present the architecture and demonstrate the major components of the system.

1 Introduction

The rapid growth of data in different disciplines has urged the development of a wide range of Internet technologies, helping to manage these data on the Web. However, there are a number of notable limitations conventional technologies are confronted with. First, building Web sites to publish large amount of data have been primarily focusing on the relational database and server side technologies [4]. This results in little reusability of software components, as most of the design is hard-coded in Web pages. Second, the only user interface for accessing the data is through static or dynamic Web pages, which are usually not personalized according to individual users preferences. Existing personalization features are usually site-specific and not reusable either. Lastly, the maintenance of Web sites is a high-effort task, due to the lack of a formalism to process the underlying data.

OntoWebber is a system to manage data on the Web with their semantics formally encoded using RDF (Resource Description Framework) [6]. It adopts an ontology-based approach, designed to overcome the aforementioned limitations for harnessing Web data. The ontologies serve as the basis for the declarative management of all types of data, so that rule-based mechanisms can reason over these data as instances of ontologies.

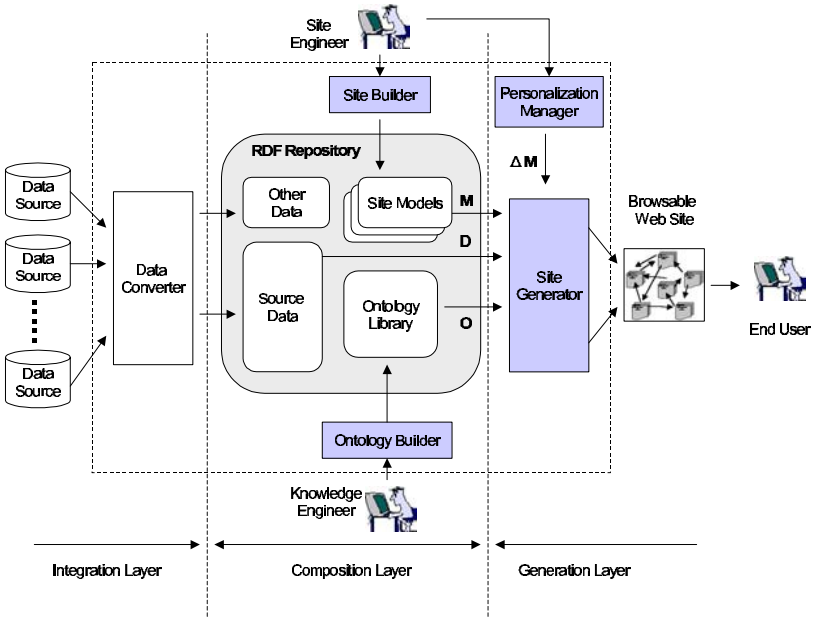

Figure 1. The OntoWebber Architecture

Proceedings of the 18th International Conference on Data Engineering (ICDE'02)
1063-6382/02 $17.00 © 2002 **IEEE**

2 Architecture

The architecture of OntoWebber system comprises three layers, namely, integration, composition, and generation services (see Figure 1). Each layer has a number of components handling specific tasks in data management.

In the integration layer, data from various sources on the Web is retrieved into local cache and converted into RDF-annotated format, based on the domain ontology. In the composition layer, ontologies are constructed and site views are created on the underlying data as site models. Finally, the browsable Web site can be generated by instantiating site models with source data in the generation layer. Personalization and verification of integrity constraints are also performed in this layer. The components of composition and generation layers will be demonstrated.

3 Demonstration

We demonstrate how to use the set of software tools in OntoWebber system to construct a portal in the Semantic Web Research Community (SWRC). The portal integrates information from various participants in a research project, and presents a centralized view tailored for individual user. The major components we will demonstrate are as follows:

(1) **Ontology Builder.** We will show how knowledge engineer can use the *ontology builder* (Protégé) to create a domain ontology and a site modeling ontology. The domain ontology to be created for this demo is specifically designed for the research community domain, whereas the site-modeling ontology is the same for any domain of interest.

(2) **Site Builder.** We will demonstrate how to publish the source data on the Web by creating site views using the *site builder*. Specifically, we will show (a) how, through a user-friendly interface, the site engineer can create a default site view (graph), which is then exported into three site models; (b) A few example rules will be defined by the site engineer using the site builder, and the execution of these rules against the site models will demonstrate the constraint verification functionality;

(3) **Site Generator**. After validation of the site view, the *site generator* will instantiate the site view while the user is interacting with the Web site. A cache is used for materializing the Web pages to reduce the response time.

(4) **Personalization Manager.** We will show how to use *personalization manager* to define model-rewriting rules, so that the site view can be modified by these rules to present personalized information access for end users. Two types of personalization will be demonstrated: (a) the user-initiated personalization allows the end user to tweak their site views through a form-based interface; (b) the system initiated personalization will update the site view according to pre-defined model-rewriting rules, so as to help end users navigating the information space.

4 Conclusion

The OntoWebber system demonstrates an ontology-based approach to site management, which was inspired by research from different areas such as database, hypermedia methodology and knowledge management [1,2,3,5,7]. It provides a set of software tools to facilitate the design, creation, generation and maintenance of Web sites. By encoding data using RDF data model as instances of a domain ontology, we can construct different site views as site models based on the pre-defined modeling ontology. Thus, site views can be tailored for individual users by re-writing the models. Maintenance of the Web site can also be automated through rule-based constraint verification against the site models.

References:

[1] Stefano Ceri, Piero Fraternali, Stefano Paraboschi: Data-Driven, One-To-One Web Site Generation for Data-Intensive Applications. VLDB 1999: 615-626

[2] Mary F. Fernandez, et al.: Declarative Specification of Web Sites with Strudel. VLDB Journal 9(1): 38-55, 2000

[3] Piero Fraternali, Paolo Paolini: A Conceptual Model and a Tool Environment for Developing More Scalable, Dynamic, and Customizable Web Applications. EDBT 1998: 421-435

[4] Susan Malaika, Resistance is Futile: The Web will assimilate your Database, Bulletin of the Technical Committee on Data Engineering, Vol. 21, No.2, pp. 4-13, 1998

[5] G. Mecca, P. Merialdo, P. Atzeni, V. Crescenzi The (Short) Araneus Guide to Web-Site Development - Second Intern. Workshop on the Web and Databases (WebDB'99) in conjunction with SIGMOD'99, May 1999

[6] Resource Description Framework (RDF). http://www.w3.org/RDF

[7] A. Mädche, S. Staab, N. Stojanovic, R. Studer, Y. Sure. SEAL - A Framework for Developing SEmantic portALs. In: BNCOD 2001 - 18th British National Conference on Databases. Oxford, UK, 9th - 11th July 2001, LNCS, Springer Verlag, 2001