Connecting *Semantic MediaWiki* to different Triple Stores Using RDF2Go

Manfred Schied[1], Anton Köstlbacher[1]. Christian Wolff[2]

University of Regensburg, [1]Information Science / [2]Media Computing
Universitätsstr. 31, 93053 Regensburg, Germany
manfred.schied@stud.uni-r.de; {anton.koestlbacher, christian.wolff}@sprachlit.uni-r.de

Abstract. This article describes a generic triple store connector for the popular *Semantic MediaWiki* software to be used with different triple stores like Jena or Sesame. Using RDF2Go as an abstraction layer it is possible to easily exchange triple stores. This ongoing work is part of the *opendrugwiki* project, a semantic wiki for distributed pharmaceutical research groups.

Keywords: triple store connector, semantic mediawiki, rdf2go

1   Introduction

Semantic MediaWiki (SMW) [1] is one of the most popular and mature semantic wiki engines currently available [2]. It is based on the MediaWiki software [3]. Queries within wiki articles allow reusing available semantic data. For extended query results or to query stored facts from outside the wiki it is necessary to connect SMW to a *triple store*.

We use Semantic MediaWiki for knowledge-based applications in the domain of pharmacy, focusing on psychiatric therapy. In the following, we briefly describe the motivation for a generic triple store connector (ch. 2), give some information on the application context (ch. 3) and explain our implementation approach (ch. 4).

2   An Abstraction Layer for Triple Stores

Triple stores are storage systems tailored for efficient storage of RDF data [4]. In addition, triple stores offer services and programming libraries for inferring new facts or for accessing data using a query language. In a distributed computing system a triple store is a rather independent component with specific features which can be utilized by an associated application programming interface (API). In addition, semantic wiki data can be used in other applications or served as linked data on the web [10].

At the moment, three different triple store products are available for use with SMW, each with a specific connector to SMW. Two of them, *RAP* [5] and *Ontoprise Basic Triplestore* [6] are based on open source software, the third one, *Ontobroker*, is

a commercial product [7]. As SMW (with *Halo Extension*, see [15]) doesn't follow W3C's recommended SPARQL/UL format exactly [8], but uses its own data format for communicating with triple stores, it is necessary to have a connector software between the two systems. To enable users of SMW to select the triple store most suitable for their needs, we have implemented a generic triple store connector using the RDF2Go library [9]. This setup abstracts from the underlying triple store and makes the storage layer easily exchangeable.

3   Application Context

The work described here has been carried out in the context of the recently started *opendrugwiki* project which itself evolved from the drug interaction database *PsiacOnline*[1]. In *PsiacOnline*, drug-interaction information in psychiatric treatment has been collected, uniformly structured, and evaluated by a team of experts in the field [11]. Transforming this approach in the direction of semantic social software appears as a logical next step: On the one hand, we expect a large community of interested experts working in psychiatry to be ready to contribute to this novel method of collecting interaction data. On the other hand, we assume that semantic wikis and the usage of structured knowledge representation standards are adequate for the given information and will allow for the answering of complex information needs.

4   Implementation Details

RDF2Go is a Java library developed at the *Forschungszentrum Informatik* (FZI) in Karlsruhe providing abstract data access methods to RDF triples stored in a triple store ("program now, decide on triple store later"[2]). It uses common adapter classes to access different triple stores. At the moment, RDF2Go delivers adapter classes for *Jena* [12] and *Sesame* [13] and can easily be extended to other triple stores. Communication between SMW and the triple store connector is done via SPARQL and the SPARUL extension [14]. Initial Loading of RDF data from SMW into the triple store is triggered with a SPARUL LOAD command on part of SMW. The connector handles this event by reading the semantic data directly from SMW's database tables due to performance reasons and a missing function for retrieving the wiki's semantic data as a whole via HTTP.

Figure 1 shows the overall architecture of our approach: The semantic media wiki accesses the triple store connector via SPARQL to retrieve query results and via SPARUL to trigger changes made in the wiki to the triple store. The triple store connector – the core component in our architecture – provides an adequate infrastructure for receiving commands and returning resulting triples using web service standards.

---

[1] *PsiacOnline* is an online service offered by *SpringerMedizin*: http://www.psiac.de
[2] Cf. http://semanticweb.org/wiki/RDF2Go and http://rdf2go.semweb4j.org/
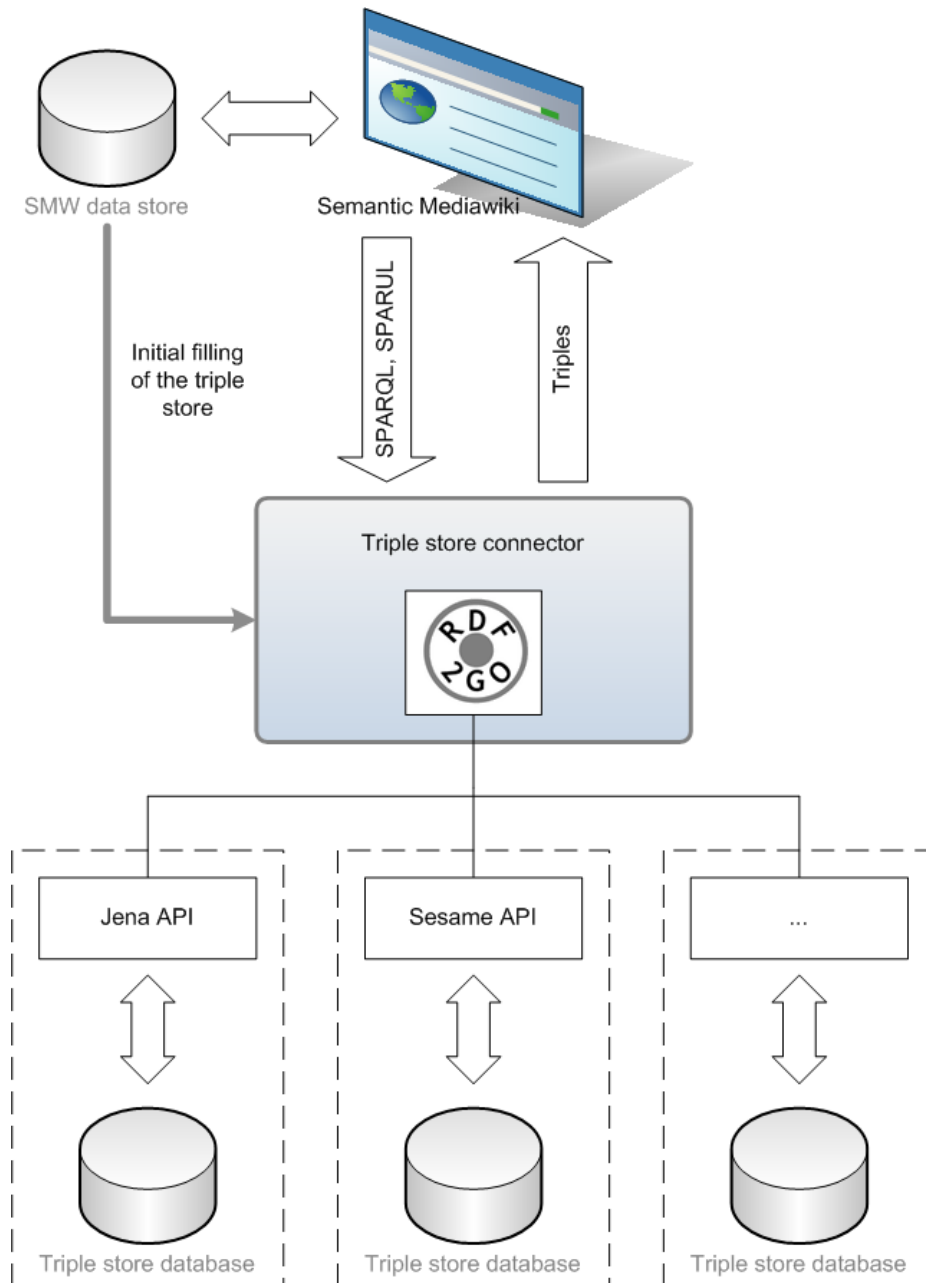
Fig. 1. SMW (with *Halo Extension* [15]) and triple store are connected via the triple store connector. RDF2Go helps to build triple store adapters for all supported triple stores.

5 Demonstration and Conclusion

For the demonstration of our approach, we will present typical usage scenarios taken from our application domain, i.e. drug interaction description and retrieval. Besides showing the feasibility of using an abstract triple store access layer, we also want to demonstrate how semantic wiki technology can facilitate search in complex structured medical data.

By making external semantic storage engines for *Semantic MediaWiki* exchangeable in an easy way, SMW can be conveniently integrated in sophisticated distributed systems. The wiki's semantic data can be re-used with other applications much better since it isn't limited to the triple store engines which have been implemented so far. This enables users of SMW to choose a triple store engine which fulfills their individual needs concerning inference and retrieval of semantic data.

6 References

1. Krötzsch, M., Vrandečić, D., Völkel, M.: Semantic MediaWiki. In: Proceedings of the 5th International Semantic Web Conference (ISWC'06). LNCS, vol. 4273, pp. 935-942. Heidelberg: Springer, (2006)
2. Köstlbacher, A., Maurus, J.: Semantische Wikis für das Wissensmanagement. In: DGI e.V./M. Heckner, C. Wolff (Ed.). Information Wissenschaft und Praxis. Mai/Juni 2009, Wiesbaden: Dinges & Frick GmbH, pp. 225-231 (2009)
3. MediaWiki contributors: Mediawiki, The Free Wiki Engine, http://www.Mediawiki.org/ w/index.php?title=Mediawiki&oldid=65192 (accessed March 3, 2010)
4. Rusher, J.: Triple Store. Position Paper, Workshop on Semantic Web Storage and Retrieval (2003). http://www.w3.org/2001/sw/Europe/events/20031113-storage/positions/rusher.html
5. Oldakowski, R., Bizer, C., Westphal, D.: RAP: RDF API for PHP. In: Proceedings of the ESWC Workshop on Scripting for the Semantic Web (2005)
6. Ontoprise GmbH (Ed.): Basic Triplestore, http://smwforum.ontoprise.com/smwforum/ index.php/Help:Basic_Triplestore (accessed March 3, 2010)
7. Ontoprise GmbH (Ed.): Ontobroker, http://www.ontoprise.de/en/home/products/ontobroker/ (accessed March 3, 2010)
8. Beckett, D., Broekstra, J.: SPARQL Query Results XML Format. W3C Recommendation 15 January 2008. http://www.w3.org/TR/rdf-sparql-XMLres/ (accessed April 21, 2010)
9. Völkel, M.: Writing the Semantic Web with Java. Technical report, DERI Galway (2005)
10. Bizer, C., Cyganiak, R., Heath, T.: How to Publish Linked Data on the Web. http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/ (accessed April 22, 2010)
11. Köstlbacher, A., Hiemke C., Haen E., Eckermann, G., Dobmaier, M., Hammwöhner R., PsiacOnline – Fachdatenbank für Arzneimittelwechselwirkungen in der psychiatrischen Pharmakotherapie. In: Osswald, Achim; Stempfhuber, Maximilian; Wolff, Christian (eds.). Open Innovation. Proc. 10 Internationales Symposium für Informationswissenschaft ISI 2007. Constance: UVK, 321-326. (2007)
12. McBride, B.: Jena: A Semantic Web Toolkit. In: IEEE Internet Computing November/December 2002, pp. 55-59. (2002)
13. Broekstra, J., Kampman, A.: Sesame: A generic architecture for storing and querying RDF and RDF Schema. Technical report. Amersfoort: Aidministrator Nederland b. v., (2001).

14. Prud'hommeaux, E., Seaborne, A.: SPARQL Query Language for RDF. W3C Recommendation 15 January 2008. http://www.w3.org/TR/rdf-sparql-query/ (accessed March 3, 2010)
15. Friedland, N.S., Allen, P.G., Matthews, G., Witbrock, M., Baxter, D., Curtis, J., Shepard, B., Miraglia, P., Angele, J., Staab, S., Moench, E. Oppermann, H., Wenke, D. Israel, D. Chaudhri, V., Porter, B., Barker, K., Fan, J. Chaw, S.Y., Yeh, P., Tecuci, D.: Project halo: towards a digital Aristotle, AI Magazine, Winter 2004, (2004), online: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.3.4804