Automated Generation of SADI Web Services for Clinical Intelligence using Ruled-Based Semantic Mappings

Mohammad Sadnan Al Manir[1], Alexandre Riazanov[2],
Harold Boley[3], Artjom Klein[2] and Christopher J.O. Baker[1,2]

[1] Department of Computer Science and Applied Statistics
University of New Brunswick, Saint John, Canada
`{sadnan.almanir,bakerc}[at]unb.ca`
[2] IPSNP Computing Inc
`{alexandre.riazanov,Artjom.Klein}[at]ipsnp.com`
[3] Faculty of Computer Science
University of New Brunswick, Fredericton, Canada
`harold.boley[at]unb.ca`

Abstract. We present a framework that automates the generation of SADI semantic web services from declarative service descriptions and semantic mappings to relational data. Mappings are specified in a Datalog sublanguage of Positional-Slotted Object-Applicative (PSOA) RuleML. We outline a novel methodology, a system architecture, and a prototype implementation for service generation. A proof-of-concept implementation and preliminary evaluation of the approach was conducted by generating a set of SADI services over a database representing a fragment of a hospital research data warehouse. These automatically generated fully functional services are used to answer simple SPARQL queries.

Infection control practitioners involved in surveillance for Hospital-Acquired Infections (HAI) need to access multiple sources of clinical data in relational databases. Ad-hoc retrieval of information from these databases requires knowledge of query languages like SQL, albeit surveillance and decision making would be more efficient if practitioners could retrieve relevant data without IT support.

Semantic Querying facilitates self-service querying by non-technical users as described recently in HAIKU [1], where SADI Semantic Web services were deployed on relational databases to provide access to HAI-related data from The Ottawa Hospital (TOH) Data Warehouse (DW). While the method successfully enables domain experts (e.g. surveillance practitioners) to semantically query relational data using terminologies from domain specific ontologies, it requires prior scripting of Java code and SQL queries to operationalize the SADI Web services. This is labor-intensive, error-prone, and mandates the availability and involvement of surveillance practitioners to consult with the IT personnel during service creation. As such, the process is a good target for automation.

Expanding on our previous work [2], we present an architecture designed to support semantic query rewriting Fig. 1, where SADI Semantic Web service code is generated automatically from declarative input and output descriptions, and a

semantic mapping of the source data in PSOA RuleML [3]. Our implementation of the architecture facilitates Web service generation without human intervention and end users are able to run queries executing the generated services using SADI query engines, such as SHARE and HYDRA [1] [4].
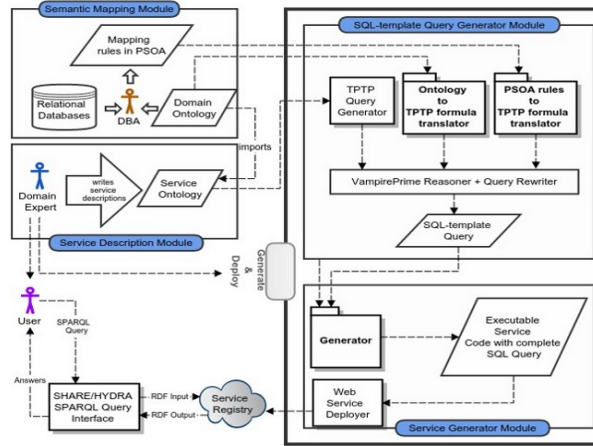

Fig. 1: Architecture supporting the automation of SADI Semantic Web service generation

The service generation process relies on four distinct modules working together: a *Semantic Mapping Module* representing the correspondence of a DB schema to a domain ontology, a *Service Description Module* where service inputs and outputs are formalized, an *SQL-template Query Generator Module*, and a *Service Generator Module* to output executable Java code for a SADI web service. Generated services are then indexed in a SADI registry.

In testing the query engine, HYDRA was able to discover the appropriate services in the registry and execute queries over the data. Based on a list of previously identified target use cases [1], our system is being further evaluated for its utilty in provisioning Semantic Querying to access HAI-related data from TOH DW. A list of frequently asked questions [5] is available for interested readers.

References

1. Riazanov A., Klein A., Shaban-Nejad A., Rose G. W., Forster A. J., Buckeridge D. L., and Baker C.J.O. Semantic querying of relational data for clinical intelligence:a semantic web services-based approach. *J. Biomedical Semantics.*
2. Al Manir M.S., Riazanov A., Boley H., and Baker C.J.O. Generating Semantic Web Services from Declarative Descriptions. In *Proc. 4th Canadian Semantic Web Symposium, co-located: Semantic Trilogy 2013, Montreal, QC, Canada.*
3. Harold Boley. PSOA RuleML: Integrated Object-Relational Data and Rules. In Wolfgang Faber and Adrian Paschke, editors, *Reasoning Web. Web Logic Rules (RuleML 2015) - 11th Int'l Summer School 2015, Berlin, Germany, July 31- August 4, 2015*, volume 9203 of *Lecture Notes in Computer Science*. Springer, 2015.

[4] http://ipsnp.com/hydra/ and http://ipsnp.co/demo-videos/
[5] http://cbakerlab.unbsj.ca:8080/swat4ls15/poster/faq.html