Using a Semantic MediaWiki to Interact with a Knowledge-Based Infrastructure

Ian Millard, Afraz Jaffri, Hugh Glaser, Benedicto Rodriguez
School of Electronics and Computer Science
University of Southampton, SO17 1BJ, UK
{icm, aoj04r, hg, br05r}@ecs.soton.ac.uk

ABSTRACT

Facilitating knowledge acquisition is a task that usually requires special purpose interfaces with which users are not familiar. Providing effective acquisition through a familiar interface, such as a wiki, can provide a route to acquiring knowledge for low user investment. We present an architecture that is being used in the ReSIST project based on a Semantic MediaWiki integrated with a knowledge base that allows users to add and view knowledge using normal Semantic MediaWiki syntax. The architecture aims to facilitate the acquisition and representation of knowledge about resilient systems for users with no experience of knowledge technologies.

Categories and Subject Descriptors

D.4.3 [**Information Systems Applications**]: Communications Applications—*Information Browsers*

General Terms

Management, Design

1. INTRODUCTION

ReSIST [3] is a Network of Excellence which integrates leading researchers from the multidisciplinary domains of dependability, security, and human factors. The general focus of this activity is the advancement and development of technologies which will ensure that future ubiquitous computing systems have the necessary properties of resilience and survivability for real world deployment. The project also aims to create architectures which are tolerant of residual development and physical faults, interaction mistakes, and malicious attacks or service disruptions.

In an effort to aid the creation of such systems, the ReSIST project has embraced the emerging principles of Ontological Engineering and the Semantic Web. This has enabled us to formally describe resilience concepts and the properties of complex components in detail, as well as information regarding people, projects and publications. Through the utilisation of these semantic representations the ReSIST project is tasked with the creation of a Resilience Knowledge Base (RKB). The RKB will combine disparate information sources with suitable user interfaces to provide a central repository which comprehensively covers all aspects of resilient computing and dependable systems. It is envisioned that the RKB will provide an invaluable resource for both researchers and students.

The RKB is intended to provide information regarding organisations that are researching resilient systems; researchers interested in resilient systems; papers associated with resilient systems; faults, errors and failures that have occurred on IST systems; and other aspects of resilient systems research topics. In addition, knowledge regarding the ReSIST project itself is recorded, including sub-project activities, meetings, work package development and management decisions.

However, the task of acquiring semantic information about an ongoing project from people who are not experts in the field of knowledge related technologies presents a significant challenge. Systems must be provided to facilitate as much incidental knowledge acquisition as possible, while still being able to gather sufficient knowledge to be meaningful to the project.

2. SEMANTIC MEDIAWIKI

A significant step in achieving incidental knowledge acquisition has been through the use and customisation of the newly developed Semantic MediaWiki (SMW) [5]. In addition to supporting general collaboration between project members, the SMW provides a means of adding metadata about the concepts and relations that are contained within the wiki. This form of 'tagging' makes it relatively simple to turn such annotations into subject, predicate, object triples that can be stored as RDF and incorporated into the RKB. Such a system has the advantage of being easy to use for non experts, but also powerful in the way in which knowledge can be created and stored.

A prototype system has been developed, utilising the SMW in conjunction with an external 3store [4] RDF repository. In the SMW, real-world or abstract entities are represented by an individual page, to which metadata can be added. The page is therefore represented as the subject resource in RDF triple form. Relations and attributes are handled differently

Posters and Demos of the 15th International Conference on Knowledge Engineering and Knowledge Management, EKAW 2006 Podebrady, Czech Republic, 2nd-6th October, 2006

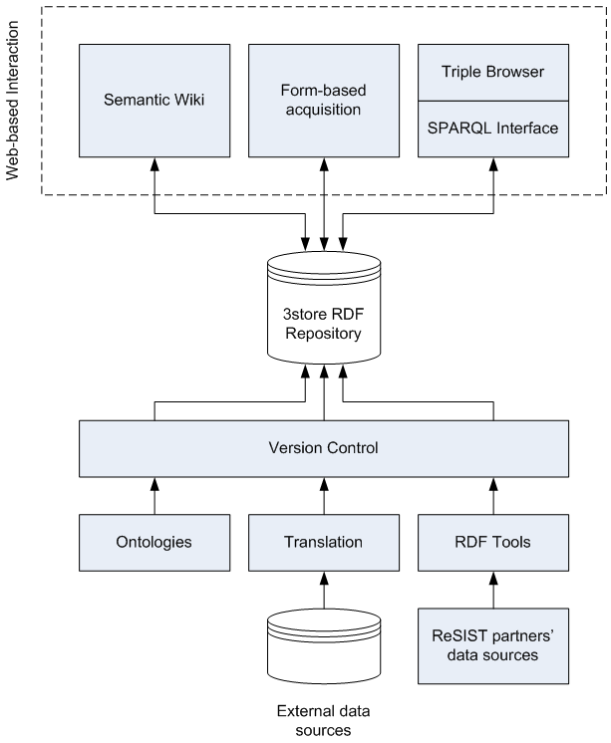

Figure 1: ReSIST RKB Architecture

within the SMW, associating the current page with another SMW resource or a literal value. When a page is saved in the SMW, custom code is invoked to export the relations and attributes as RDF and to assert them within the RKB repository.

For example, a page in the SMW may describe a publication, to which a project member wishes to be associated as an author. Utilising the AKT Ontology [1], the following relation may be inserted into the page, specifying the desired fact.

```
[[has author::User:Joe Bloggs|Joe Bloggs]]
```

Entering this special SMW markup within the page will cause the following triple to be asserted into the RKB:

```
<http://resist.eu/publications/Bloggs06>
<http://www.aktors.org/ontology/portal#has-author>
<http://resist.eu/people/Joe-Bloggs> .
```

However, one area in which the SMW lacks good support is that of namespaces, as general facilities for utilising external ontologies, concepts and data-types are yet to be implemented. In a closed world SMW deployment this is not a problem, and indeed simplifies the input required by users. Nevertheless, namespaces are vital for disambiguation and ontological inference, so the export routines apply the relevant namespace prefixes during RDF generation. This is achieved by using a static mapping between the SMW representation of ontological concepts and their external 'real-world' form in the RKB.

The SMW can also be used as a means of exposing knowledge stored within the RKB. For example, pages describing the classes and properties from external ontologies have been imported into the SMW, permitting users to view and discuss the rationale behind each. As well as facilitating collaborative ontology development, these representations allow users to readily see whether relations and concepts have been used appropriately when entering semantic markup.

In addition to knowledge obtained through the use of the SMW, significant efforts are underway to facilitate the acquisition of semantic metadata from external sources. Tools have been developed to allow non-expert users at each of the 18 ReSIST partner sites to periodically generate RDF data regarding their institution and its activities. This information is then 'pushed' to the RKB server through a version control mechanism and automatically asserted, maintaining an up-to-date representation of the disparate information sources. Work has also been done to allow the bulk-import of information from Cordis, the ACM publications database, Citeseer and smaller-scale EPrints repositories.

Finally, a generic web-based form interface has been developed which can be configured to allow the acquisition of information into a specific ontologically mediated format. This interface is currently being used to collate user-submitted data regarding university courses taught to students that are related to various aspects of resilient and dependable systems.

3. FUTURE WORK

The ReSIST project is currently in its ninth month and the benefits of using the RKB architecture can already be seen. Content acquisition will be an ongoing process, and should enable more interesting analysis to be performed once a more substantial data-set is available. Combined with this effort is a requirement to develop interfaces with which novice users can easily explore the RKB, which may potentially include extensions of the work demonstrated in the CS AKTive Space project [2]. However, the maintenance of large semantic data sets presents its own challenges, not least of which are issues regarding referential integrity of knowledge acquired from multiple sources.

4. ACKNOWLEDGMENTS

This work is supported under the ReSIST Network of Excellence, which is sponsored by the Information Society Technology (IST) priority in the EU Sixth Framework Programme (FP6) under contract number IST 4 026764 NOE.

5. REFERENCES

[1] AKT Ontology. http://www.aktors.org/ontology/.

[2] CS AKTive Space. http://cs.aktivespace.org/.

[3] The ReSIST Project. http://www.resist-noe.org/.

[4] S. Harris and N. Gibbins. 3Store: Efficient bulk RDF storage. In *Proceedings of the 1st International Workshop on Practical and Scalable Semantic Systems*, pages 1–15, 2003.

[5] Semantic MediaWiki. http://meta.wikimedia.org/wiki/Semantic_MediaWiki.