CEUR Workshop Proceedings
http://ceur-ws.org
ISSN 1613-0073

Vol-596
urn:nbn:de:0074-596-3

ORES-2010
Ontology Repositories and Editors for the Semantic Web

Proceedings of the 1st Workshop on Ontology Repositories and Editors for the Semantic Web

Hersonissos, Crete, Greece, May 31st, 2010.

Edited by

Mathieu d'Aquin, **The Open University, UK**
Alexander García Castro, **Universität Bremen, Germany**
Christoph Lange, **Jacobs University Bremen, Germany**
Kim Viljanen, **Aalto University, Helsinki, Finland**

10-Jun-2010: submitted by Christoph Lange
11-Jun-2010: published on CEUR-WS.org

Semantic Interoperability Framework for Estonian Public Sector's E-Services Integration*

Kalle Tomingas[1], Martin Luts[2,3]

[1]Department of Computer Science, Tallinn University of Technology, Estonia,

[2]Department of Informatics, Tallinn University of Technology, Estonia

[3]ELIKO Competence Centre in Electronics-, Info- and Communication Technologies, Tallinn, Estonia

June 10, 2010

Abstract

E-services, based on automated data exchange in distributed technological and organizational environment, are an effective way to build cross-border, controlled information services. Processes of creation, integration, management, reuse, discovery and composition of e-services are not very efficient without understanding the meaning of information resources. Creation and management of human and machine readable semantics of heterogeneous and distributed information resources are more complicated than coordinated documentation process, and require new interoperability principles, architecture and infrastructure. This paper outlines the idea and architecture of the Estonian semantic interoperability initiative in the public sector. The paper presents a collaborative ontology engineering toolset and repository as a part of interoperability infrastructure, built with Semantic Mediawiki, to manage the semantics of information resources.

Keywords: semantic interoperability, ontology, web service, web service annotation, semantic mediawiki, rdf/rdfs, owl, wsdl/sa-wsdl.

1 Introduction

This paper describes the Estonian public sector's semantic interoperability initiative and outlines the architecture of semantic integration of State Information Systems and e-services. The paper provides a framework, a toolset and a collaborative semantic assets management solution in the Semantic Mediawiki[1] (SMW) environment. We concentrate on the practical implementation of infoware metadata management like web services descriptions, ontologies, classifiers, their creation, import and export, versioning, collaboration and change management. The paper shows how annotation and linking of different assets allows us to realize the semantic interoperability idea, manage the heterogeneous information resources and give them human and machine readable semantics. The given semantics is used for search, mange and re-use of existing information assets components, quality insurance, new web services discovery and composition.

Examples of semantic interoperability initiatives can be found in different EU countries: UK Public Sector Linked Data Government [2], Germany's Deutschland Online[3],

*This research was supported by European Social Fund's Doctoral Studies and Internationalisation Programme DoRa

[1]http://semantic-mediawiki.org

[2]http://data.gov.uk/

[3]http://www.deutschland-online.de/DOL_en_Internet/broker.jsp

Italian Public Administration[4], Finnish FinnONTO[5] [2], Estonian Semantic Interoperability Framework [1] or Latvian's Semantic Latvia [3]. Ideas behind those projects vary from human readable semantic descriptions to a large-scale international semantic interoperability infrastructure. Other examples about Pan-European semantic interoperability initiatives are SEMIC (SEMantic Interoperability Centre Europe)[6], led by the European Commission's ISA[7] program and SemanticGov[8], which is targeted to provide integrated public services to citizens at the national or Pan-European level [4]. SEMIC is designed as a brokerage platform for existing third party semantic assets (e.g. classification lists, ontologies, etc.). The aim of SemanticGov is to build the new infrastructure (software, models, services, etc.) necessary for enabling the offering of semantic web services for public administration agencies, within and between the EU countries.

2 Semantic Interoperability Architecture

The semantic interoperability architecture for state information system and registries of Estonia (see Figure 1) consists of the following interrelated components: ontologies and semantically annotated objects supported by policies and guidelines, several processes and workflows, tools, educational activities, PR among others. The Administration System for the State Information System [5]and Semantic Asset Management Environment (SEHKE) are the central tools in the semantic interoperability architecture for the state information system. SEHKE fulfills the following tasks in the semantic interoperability architecture: hosting and publishing of ontologies, infoware's metadata including semantic annotations; serving as a semantic search engine for semantic assets (resources).

SEHKE works to ensure the interoperability of public sector registries and the reuse of technical, organizational and semantic resources. For end-users, SEHKE is the tool for obtaining information about existing services (as well as about service descriptions and the principles of service provision) and apply for the right to use a service or propose the creation of a new service.

One of the main components of the semantic interoperability architecture – as designed for Estonian state registries – is ontology. We use the term 'ontology' meaning "a formal explicit specification of a shared conceptualization for a domain of interest" [6] and "information about how individuals are grouped and fit together in a particular domain". Currently, there are some ontologies in Estonian, which could be used to annotate operations performed by state registries and other objects. The ontology component in our architectural framework is not a monolithic structure – for the purposes of easier, domain-expert driven maintenance, it is divided into subject areas or domains, e.g. "Environment", "Social Affairs". The initial tree for categorizing ontologies is based on the official naming of EU activities. The language used in the semantic description of ontology objects draws from W3C recommendation OWL (Web Ontology Language)[9]. Ontologies are developed and maintained in a distributed manner, but stored and published centrally. The tools to be used for ontology maintenance are not prescribed as long as certain standards are followed (e.g. Collaborative Protégé, Semantic Mediawiki or others). In the next section, we give an overview of a new online collaborative tool for semantic resources management, which is playing an important

[4] http://www.cnipa.gov.it/site/it-IT/
[5] http://www.seco.tkk.fi/projects/finnonto/
[6] http://semic.eu/about_semantic_interoperability_centre_europe.html
[7] http://ec.europa.eu/isa
[8] http://www.semantic-gov.org/
[9] http://www.w3.org/TR/2004/REC-owl-features-20040210

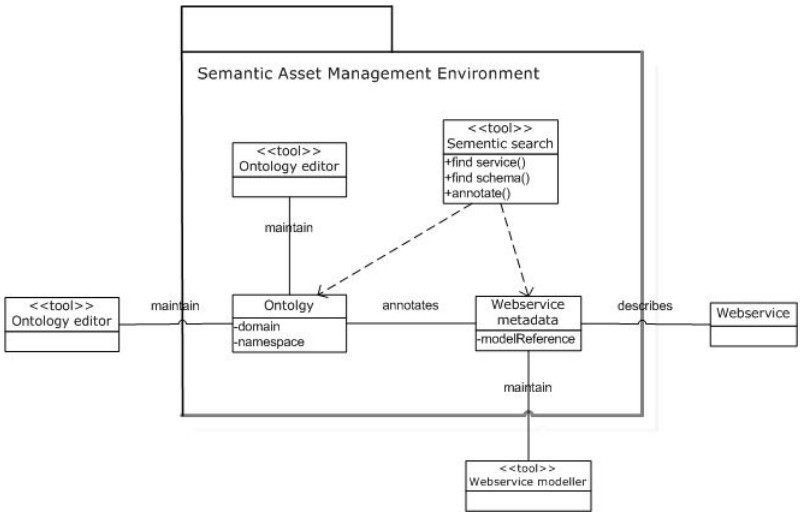

Figure 1: Semantic interoperability conceptual architecture for state information system and registries

part in the large-scale semantic interoperability initiative.

3 Collaborative Semantic Resource Management

SEHKE is a web-based collaborative environment for creating, management and annotation of semantic resources and assets of state information systems and registries.. SEHKE is built to handle information resources like ontologies, concepts and properties, subject areas (domains of interests), web service descriptions and classifications. SEHKE environment is one piece in the big picture of large-scale semantic interoperability architecture. The functionality of the SEHKE toolset covers the creation process of the information assets (e.g. ontologies, web service descriptions, classifiers), the semantic annotation of web service descriptions, collaboration, responsibility, notifications and approval, quality and version control, import and export, semantic search and finally (re)usage of all assets. The SEHKE environment is set up and implemented in the Semantic Mediawiki environment, where the needed functionality is built with semantic extensions, templates and forms, existing plugins and new developed components for special requirements (import and export).

System concepts of the SEHKE environment are based on the main categories of semantic assets. System concepts are implemented as wiki pages, with certain types of mandatory properties and a set of defined forms, to create and manage a content of different concept types. One required property is "Type", which is the main instrument to categorize concepts (e.g. `[[Type::Ontology]]`). Each SEHKE concept has its own set of required or allowed properties that are used to implement the possible relationships between different system concepts (e.g. ontology page has property `[[SubjectArea::Economy]]` to define a relation that the ontology belongs to the subject area). Named and implemented properties and their values form the schema of concepts (Figure 2) that represents the semantic interoperability requirements and the structure of SEHKE functionality.

Ontology is one of the key concepts of the SEHKE tool and it stands for formalized semantics. Ontology is a named set of concepts, their definitions, hierarchies and properties that represents the conceptualization of the domain or subject area, which

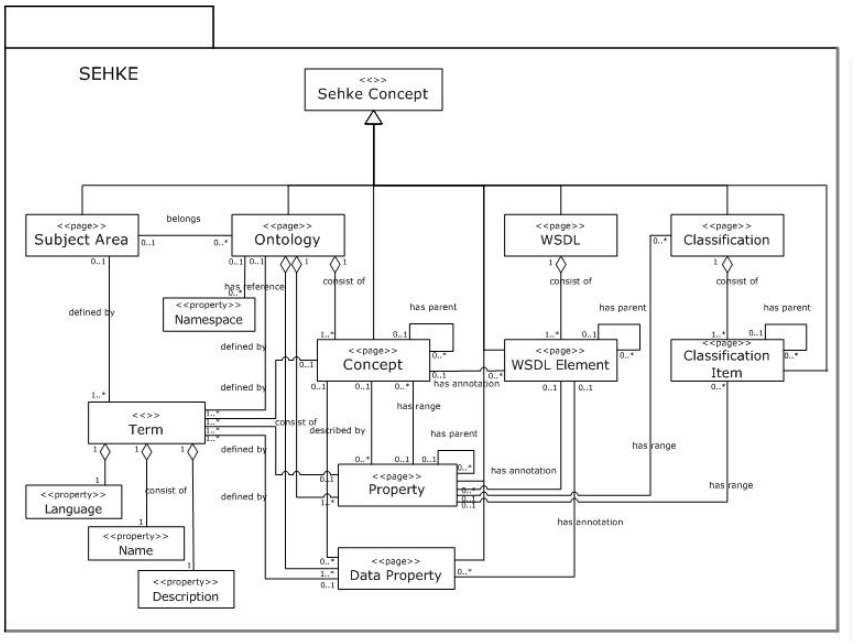

Figure 2: Schema of system concepts of SEHKE environment

is in the interest of the Estonian public sector. Ontology and its concepts are besides web services descriptions and classifications a main type of semantic assets. Ontology belongs to the subject area that frames the domain and the responsibility of the owner of the ontology. Ontology is implemented as a special page, with the idea to group and collect related classes, their properties, instances and rules. Ontology, its member concepts and properties (data and object properties) can have multiple terms (set of names and descriptions) in different languages, which explain the meaning of the item. Concepts (classes) are organized into taxonomies or hierarchies (e.g. using property definition `[[subclassOf::ParentConceptName]]` to associate concept page with a higher level concept), and they can have certain formalization rules that can be automatically checked and validated by automatic reasoning programs. The creation of ontologies, concepts and properties is simplified through automated detection of similar assets, supported by syntactic and semantic search algorithms and backed with the different language techniques. Ontology creation and management has online collaborative support functions, like email notifications, change acceptance and an iterative process for achieving consensus between multiple parties. The collection of all SEHKE pages forms the general wiki namespace (e.g. http://www.sehke.eu/ is the namespace of public instance of the SEHKE environment). In addition to the wiki's namespace, ontology has special "Namespace" property (e.g.), for different namespace values that are used in ontology export/import to rdf/owl. The multiuser collaborative editing, multilingual terms and descriptions and configurable language techniques makes the SEHKE a real multi-national semantic tool.

WSDL is a SEHKE concept for web services descriptions that documents e-services of the Estonian state information system's data transportation layer X-Road (Estonian X-Tee)[10]. X-Road currently supports the technical interoperability of the components of state information systems and registries and SEHKE is like an extension for adding a semantic layer top of the X-Road infrastructure. Web services description documents

[10]http://www.riso.ee/en/information-policy/projects/x-road

are semantic assets, which can be imported and exported in WSDL/SA-WSDL[11] language. The essential idea of semantic interoperability is to add semantic annotations to web services input/output data elements and structures and related operations. Defined ontology's concepts and data or object properties are the key ideas that are used for WSDL annotations. An annotation is a semantic relationship between a WSDL element and an ontology's concept, which can be created manually or semi-automatically (in the future releases), using syntactic and semantic search algorithms, to match similar terms with the concepts (e.g. SAWSDL *modelReference* extension attribute used to associate WSDL elements with a semantic model). Semantic annotation of web services brings us to the next level, which makes available automatic discovery of web services and composition of new web services. The composition of new services is based on existing services, operations and their input/output data elements, which have a semantic match on the conceptual level.

4 Conclusions

The main approach to extend semantic interoperability in communication and information exchange process is to build a higher semantic integration layer on top of participating information systems and organizations. Mapping critical communication structures against ontologies and the same semantic concepts using annotation techniques adds human- and machine-readable semantics to the process. Opening the meaning of system components and functions, integrating computer programs, linguistic techniques and human touch to knowledge management process, we fulfill the promises and expectations of semantic interoperability idea and creating better public sectors services. The SEHKE toolset that is powered with semantic functionality of the SMW is the open platform for collaboration and semantic asset management (e.g. ontologies) and meant to be part of federated pan-European repository initiative: SEMIC.EU [7].

References

[1] Haav, H.M., Kalja, A., Küngas, P., Luts, M.: Ensuring Large-Scale Semantic Interoperability: The Estonian Public Sector's Case Study. Databases and Information Systems V, IOS Press 2009, pp.117-129.

[2] Hyvönen, E., Viljanen, K., Tuominen, J., Seppälä, K.: Building a National Semantic Web Ontology and Ontology Service Infrastructure - The FinnONTO Approach. In Proceedings of the European Semantic Web Conference (ESWC 2008), Tenerife, Spain, 2008, (available at http://www.seco.tkk.fi/publications/2008/hyvonen-et-al-building-2008.pdf)

[3] Barzdins, J., Barzdins, G., Balodis, R., Cerans, K., Kalnins, A., Opmanis, M., Podnieks, K.: Towards Semantic Latvia. In Proceedings of Communications of 7th International Baltic Conference on Databases and Information Systems. Vilnius, Lithuania, 2006, pp. 203-218, (available at http://melnais.mii.lu.lv/audris/CReady_Barzdins_TowardsSemanticLatvia.pdf)

[4] Vitvar, T.: Infrastructure for the Semantic Pan-European E-government Services. In Proceedings of the 2006 AAAI Spring Symposium on The Semantic Web meets e-government (SWEG), 2006, (available at http://www.semantic-gov.org/index.php?name=Web_Links&req=visit&lid=4)

[5] Vallner, U.: Nationwide Components of Estonia's State Information System. In Baltic IT&T Review 3(42), 2006.

[6] Gruber, T.R.: A Translation Approach to Portable Ontology Specifications. In Knowledge Acquisition, Vol. 5, 1993, pp.199-220.

[7] Reichling, K., Luts, M., Fahl-Spiewack, R.: A pan-European repository: SEMIC.EU as the point of reference for eGovernment ontologies, ORES, 2010

[11] http://www.w3.org/TR/2007/REC-sawsdl-20070828