SWING: An Integrated Environment for Geospatial Semantic Web Services

Mihai Andrei[1], Arne Berre[2], Luis Costa[2], Philippe Duchesne[3], Daniel Fitzner[4],
Miha Grcar[5], Jörg Hoffmann[6], Eva Klien[4], Joel Langlois[7], Andreas Limyr[2],
Patrick Maue[4], Sven Schade[4], Nathalie Steinmetz[6], Francois Tertre[7],
Laurentiu Vasiliu[1], Raluca Zaharia[1], and Nicolas Zastavni[3]

[1] DERI Galway, Ireland
⟨first⟩.⟨last⟩@deri.org
[2] SINTEF Group, Oslo, Norway
⟨first⟩.⟨last⟩@sintef.no
[3] IONIC Software, Liege, Belgium
⟨first⟩.⟨last⟩@ionicsoft.com
[4] University of Münster, Germany
⟨first⟩.⟨last⟩@uni-muenster.de
[5] Jozef Stefan Institute, Ljubljana, Slovenia miha.grcar@ijs.si
[6] STI Innsbruck, Austria
⟨first⟩.⟨last⟩@sti2.at
[7] BRGM, Orleans, France
⟨first-initial⟩.⟨last⟩@brgm.fr

Abstract. Geospatial Web services allow to access and to process Geospatial data. Despite significant standardisation efforts, severe heterogeneity and interoperability problems remain. The SWING environment[1] leverages the Semantic Web Services (SWS) paradigm to address these problems. The environment supports the entire life-cycle of Geospatial SWS. To this end, it integrates a genuine end-user tool, a tool for developers of new Geospatial Web services, a commercial service Catalogue, the Web Service Execution Environment platform (WSMX), as well as an annotation tool. The demonstration includes three usage scenarios of increasing complexity, involving the semantic annotation of a legacy service, the semantic discovery of a Geospatial SWS, as well as the composition of a new Geospatial SWS.

1 Introduction

Geospatial Web services provide access to, and processing functions for, Geospatial data. The need for sharing and processing such data on a large scale has lead to significant standardisation efforts by the *Open Geospatial Consortium (OGC)* and the Technical Committee 211 of the International Organisation for Standardisation (ISO/TC 211). The OGC drives standardisation efforts regarding Geospatial languages such as the Geographic Markup Language (GML)[1], GeoRSS, or the Keyhole Markup Language (KML)[2]. Also, standardised interfaces for certain types of Geospatial Web services have been

[1] Developed in the SWING project (Semantic Web services INteroperability for Geospatial decision making), funded by the European Commission (FP6-26514).

S. Bechhofer et al.(Eds.): ESWC 2008, LNCS 5021, pp. 767–771, 2008.

fixed, such as the *Web Feature Service (WFS)* which provides access to a Geospatial database. However, these specifications are purely syntactic, and so severe heterogeneity and interoperability problems remain, with little support for resolving them. The WFS protocol is for example syntactic in the way that the optional meta-information that can be provided to describe the geo-information (features) is not formalized. The SWING environment leverages the *Semantic Web Services (SWS)* paradigm to address these problems.

The SWING environment supports the entire life-cycle of Geospatial SWS; to this end, it integrates: a genuine end-user tool (called *MiMS*); a tool (*DEV*) supporting the composition of new Geospatial Web services from existing ones; a commercial service Catalogue (*CAT*); a state-of-the-art SWS platform (*WSMX*); and a tool (*ANNOT*) that helps with the creation of the semantic annotations. Interconnecting all these disparate components is a challenge in itself. Our solution shows how all the issues regarding the required interplays can be resolved, and hence how Geospatial SWS can be realized. Also, various contributions are made regarding base technologies, such as Geospatial SWS ontologies, and particular methods for semantic annotation and discovery.

Section 2 overviews the SWING environment and outlines our technical contributions. Section 3 explains what will be demonstrated at ESWC'08. Section 4 wraps up.

2 The SWING Environment

Figure 1 illustrates the SWING environment, in terms of its components and their intuitive relations [3]. MiMS is the environment for the Geospatial domain expert, i.e., the end-user who will need access to Geospatial data; MiMS is a genuine tool used at the French Geological Survey (BRGM), the leading institution in France for Geospatial decision making. DEV is a UML modelling tool adapted for the Geospatial domain, enabling human developers to conveniently compose new Geospatial Web services from existing ones. Both MiMS and DEV require the ability to discover existing Geospatial SWS, and to semantically annotate services (legacy services in MiMS, new composed services in DEV) for later use. These functionalities are supported by the CAT, WSMX, and ANNOT components; they interface to MiMS and DEV via the query annotation GUI (*Query Annot*) for discovery queries, and via the service annotation GUI (*Service Annot*) for annotating services. Query Annot accesses ANNOT for helping to construct the semantic part of the discovery query; Service Annot does the same for the semantic annotation. ANNOT is a tool developed especially for the SWING environment, employing term matching techniques and ontology structure analysis to help map a natural language description of a query/service into a semantic query/annotation. CAT is an adaptation of a commercial Geospatial Web service Catalogue, handling service storage and discovery. For executing a discovery query, CAT performs the spatial part of the query itself, and then calls WSMX [4] for semantic matching. New Geospatial Web services composed in DEV are automatically exported into the orchestration format understood by WSMX; these orchestrations are registered and discovered in CAT just like any other Geospatial SWS; when executing an orchestration, the *WFS Wrapper* component sits between WSMX and MiMS so that, from the end-user point of view, the composed service exposes a standard OGC interface. In this way, WSMX technology can be seamlessly integrated into existing Geospatial Web service environments.

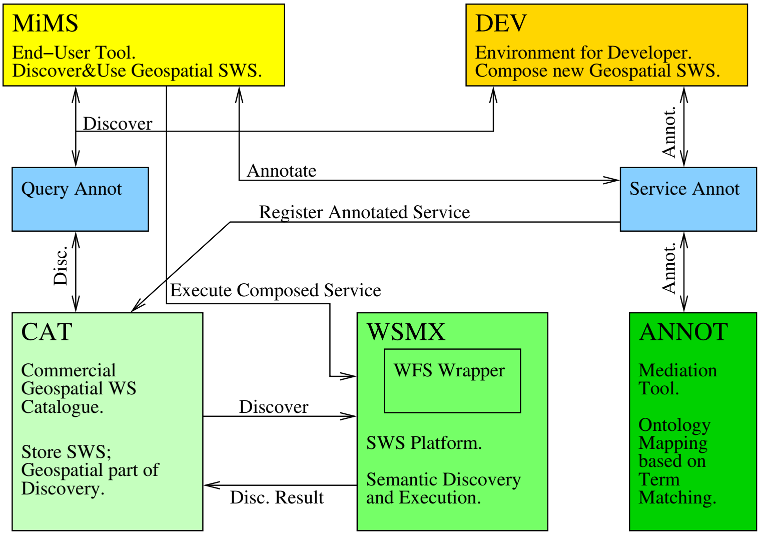

Fig. 1. A schematic overview of the SWING environment

It should be clear that the integration of such a diverse set of components involves significant engineering issues. Also, the construction of such a complex environment involves numerous design decisions, which we cannot describe in detail for lack of space.

Apart from the methodological and software contributions made by the SWING environment, many of its elements are technical contributions in themselves; in particular:

- **Geospatial ontologies.** We created WSML [5] ontologies formalising many aspects of (the exploitation of) mineral resources. The necessary domain knowledge was acquired from domain experts at BRGM [6].
- **Semantic annotation of WFS and WPS.** We designed new strategies to annotate and discover WFS and *Web Processing Services (WPS)*; In difference to previous approaches [7,8] our strategies are light-weight in that they use query containment in Datalog (in WSML-Flight [5], to be precise) for the required matching and keep the annotations simple. In addition, we designed techniques to automatically suggest semantic annotations by mapping natural language statements into our ontologies [9]. A GUI supports the selection of a corresponding ontology sub-graph. It is automatically transformed into the required WSML-Flight expression.
- **Enhancement of geospatial catalogue discovery capabilities.** The catalogue interface is used in an alternate way by exchanging the keyword set for the thematic search with more structured WSML goals [10]. Furthermore the functionality of the catalogue is enhanced by delegating the thematic search to a WSMX server.
- **Convenient UML modelling of composed Geospatial Web services.** DEV allows to orchestrate existing Web services into new ones at UML level [3]. Support is provided, e.g., for integration issues regarding WFS: being an entry point into a database, every WFS provides a rich set of possible outputs, with non-uniform structure and naming. We resolve this using semantic annotations.

- **Exporting UML into semantic orchestrations.** WSMX supports semantic orchestrations by *Abstract State Machines (ASM)*, which operate on semantic data and are very flexible in orchestration specification and execution. We designed an automatic transformation from UML orchestrations into ASMs.
- **Executing semantic orchestrations.** ASM execution is a very powerful and flexible mechanism which relies heavily on reasoning and hence pays a computational prize. By a number of enhancements to ASM execution in WSMX, we obtained speed-ups of several orders of magnitude.

3 Demonstration

The demonstration focuses on the “big picture” of the integrated environment, showing how users interact with it, in three usage scenarios of increasing complexity; technical issues are highlighted on the way. The first scenario shows how a MiMS user annotates a legacy Geospatial Web service via an interface to ANNOT, and how the service is registered in CAT. As an example, a WFS that provides information on quarry boundaries and aggregate production rates of quarries is annotated. The second scenario shows how a MiMS user creates a semantic discovery query supported by ANNOT, and how that query is executed in an interplay of CAT and WSMX. In this step, the previously annotated quarry WSF is discovered. The third scenario shows how a DEV user composes a new WFS from existing Web services, with a graphical modelling tool; the new WFS is annotated, registered, and exported into an ASM for execution by WSMX; the MiMS user discovers the new WFS, and the WSMX execution is hidden behind a standard OGC interface. The quarry WFS is used as part of the composition.

Consider the third and most complex scenario. A MiMS user needs a WFS providing aggregated mineral resource consumption/production data for a particular area in France [11]. This is a classical business use case including complex systems with complex Geospatial data models. No WFS can be found that delivers the desired data; so a developer in DEV must create such a WFS. The developer uses Query Annot, the semantic discovery interface, to discover several WFS delivering either consumption or production data. The previously mentioned quarry WFS is used to serve production data. In each discovery step, a short natural language statement suffices to select concepts and relations from the domain ontology, from which the required WSML-flight query is automatically created. After discovering the services, the developer combines them. A particular detail highlighted by the demonstration is the following. A major source of complications when integrating several WFS is the need to integrate data from diverse WFS with non-uniform output structure and naming. This issue is addressed by means of semantic annotations: the DEV user identifies, only once,the desired domain ontology concept; the correct output of each WFS is then automatically selected based on its annotation. For non-annotated WFS, a term matching tool can directly be used to find the output attribute that is the best match for the desired concept.

Once the graphical model is complete, the developer pushes a button to export it into an ASM. The service is then annotated using Service Annot, selecting concepts and relations from the domain ontology similarly as in the discovery interaction explained above. Now the new WFS is discovered in MiMS, and the MiMS user invokes it just like any other WFS. The invocation is handled by the WFS Wrapper component. Towards MiMS, this acts like a standard OGC interface; it communicates to WSMX the parameter values needed for the ASM execution, and it extracts from the outcome of the execution the data desired by MiMS. The demonstration shows the WFS Wrapper at work and, if desired, the trace of the ASM execution that is behind it.

4 Summary

The SWING environment comprises support for the entire life-cycle of Geospatial semantic Web services. From a Semantic Web perspective, this is a detailed case study of putting semantics to use; the case study is important due to the omni-presence of Geospatial data in human society. The SWING environment shows how the diverse involved components can be integrated. This is best presented, by far, in the form of a system demonstration illustrating the use of the environment as a whole.

In contrast to previous affords in connecting Geospatial Web services with Semantic Web technology like [12,13], SWING aims at an integrated framework for annotation, development, discovery and execution of Geospatial SWS. The demonstration represents a related best practice. A second major difference to other approaches is the use of WSML, specifically of the variant WSML-Flight. WSML-Flight offers alternate ways for service annotation, ontology formalisation and discovery

References

1. OGC: Open Geospatial Consortium, Inc., Geography Markup Language (GML) Encoding Specification (GML) Version 2.1.2 (2002)
2. OGC: Open Geospatial Consortium, Inc., KML 2.2: An OGC Best Practice. Best Practice Paper (2007)
3. Hoff, H., et al.: D6.1 The Architecture of the Development Environment. Deliverable of the SWING Project (2006)
4. Fensel, D., et al.: Enabling Semantic Web Services: The Web Service Modeling Ontology. Springer, Heidelberg (2006)
5. de Bruijn, J., Lausen, H., Polleres, A., Fensel, D.: The web service modeling language: An overview. In: Sure, Y., Domingue, J. (eds.) ESWC 2006. LNCS, vol. 4011, pp. 590–604. Springer, Heidelberg (2006)
6. Klien, E., Schade, S., Hoffmann, J.: D3.1 Ontologies in the SWING Application - Requirement Specification. Deliverable of the SWING Project (2007)
7. Lutz, M., Klien, E.: Ontology-based retrieval of geographic information. International Journal of Geographic Information Science 20(3), 233–260 (2005)
8. Lutz, M.: Ontology-based descriptions for semantic discovery and composition of geoprocessing services. Geoinformatica (2006)
9. Grcar, M., et al.: D4.1 Representational language for Web-service annotation models. Deliverable of the SWING Project (2006)
10. Duchesne, P., Zastavni, N.: D5.1 Online services. Deliverable of the SWING Project (2007)
11. Langlois, J., et al.: D1.1 Use Case Definition and I&T Requirements. Deliverable of the SWING Project (2007)
12. Klopfer, M., Kanellopoulos, I.: Orchestra - an open service architecture for risk management (2008)
13. Lemmens, R.: Semantic interoperability of distributed geo - services. PhD thesis, Nederlandse Commissie voor Geodesie (NCG), Delft (2006)