FRONT PAGE

PROJECT FINAL REPORT

Final Publishable Summary Report

Grant Agreement number: INFSO-ICT-215876

Project acronym: SERVICE-FINDER - REALIZING WEB SERVICE DISCOVERY AT WEB SCALE

Project title: Service-Finder:

Funding Scheme: STREP-D

Period covered: from January 2008 to December 2009

Name of the scientific representative of the project's: Prof. Emanuele Della Valle

Organisation: CEFRIEL

Tel: +39 02 23954 324

Fax: +39 02 23954 505

E-mai: pm.service-finder@cefriel.it

Project website address: http://www.service-finder.eu

Final Publishable Summary Report

http://www.service-finder.eu http://demo.service-finder.eu

Service Oriented Architecture (SOA), implemented using WSDL and UDDI, reached in 2006 the plateau of productivity[1]. Several success stories have been told about the benefit of adopting a SOA inside the organizational boundaries. Using McKinsey slogan, SOA enables "flexible IT for better strategy"[2], which means both less maintenance costs and shorter time to market for new IT enabled products. A key ingredient for achieving such flexibility is service discovery.

Searching for Web Services on the Web today is a hard task. Public UDDI registries are not available; general-purpose search engines are not precise enough; and dedicated service portals cover only a small fraction of the services published on the Web. Moreover, useful information about services such as documentation, price, terms of usage is scattered through different documents on the Web.

Early attempts to use UDDI to organize such information were unsuccessful. The UDDI Business Registry was shut down in 12 January 2006[3]. The Semantic Web community, at that time, believed that the reason for such a failure was the lack of methods and models to describe and match services in a heterogeneous and open-end environment such as the Web. OWL-S[4], WSMO[5], and SAWSDL[6] were developed. However all of them are still lacking adoption. Community maintained portals such as xMethods 2 and Programmable Web 3 emerged as a dominant solutions instead. Apparently, service descriptions provided by humans for humans outperformed any attempt to provide machine processable descriptions, but a closer look to such portals tells that even them document only a small fraction 4 of the services available on the Web, less than 1%.

The great challenge of bringing SOA out of the organisational boundaries is automatically making sense of all information about services that providers publish on the Web in addition to WSDLs.

Service Finder embraced the great challenge of automatically making sense of all information about services that providers publish on the Web in addition to WSDLs. Such additional information

---

[1] "Hype cycle special report evaluates maturity of 1,650 technologies," Gartner, Tech. Rep., July 2009.

[2] "Flexible it, better strategy," McKinsey, Tech. Rep., November 2003, http://www.mckinseyquarterly.com/Business Technology/Application Management/ Flexible IT better strategy 1346.

[3] "Uddi business registry update," OASIS, Tech. Rep., December 2005.

[4] D. Martin, M. Burstein, J. Hobbs, O. Lassila, D. McDermott, S. McIlraith, S. Narayanan, M. Paolucci, B. Parsia, T. Payne, E. Sirin, N. Srinivasan, and K. Sycara, "OWL-S: Semantic markup for web services," November 2004, w3C Member Submission, available from http://www.w3.org/Submission/2004/07/.

[5] H. Lausen, A. Polleres, and D. Roman, "Web service modeling ontology (WSMO)," World Wide Web Consortium, Submission, June 2005. [Online]. Available: http://www.w3.org/Submission/2005/SUBM-WSMO-20050603

[6] J. Farrell and H. Lausen, "Semantic Annotations for WSDL and XML Schema," http://www.w3.org/TR/sawsdl/, August 2007.

consists of web pages that, for instance, document the service, present pricing policies, explain terms and condition. To do so, we developed a Semantic Search Engine for Web Services in which Web Services are embedded in a Web 2.0 environment. On the one hand, Service-Finder is a search engine that follows the standard architecture of search engines. It has a crawler, an indexer, a ranker and a search interface. In addition to these components, it has an ontology and an automatic annotator, which qualify it as a Semantic Search engine. On the other hand, Service Finder is a Web 2.0 portal were users are invited to add tags and comments to services, and to categorize and rate services. All this user generated content is used to train the Automatic Annotator and to identify clusters of related services.

The Beta version of the Service-Finder portal is live since October 2009 at http://demo.service-finder.eu. Web users are able to test its functionalities and provide their contributions following a collaborative approach typical of Web 2.0 sites. Even if the project is approach its end, partners will continue to keep the portal up and running until 2012.

The main components

The Service-Finder Ontology plays a central role in our concept; all Service Finder components either populate this ontology with instances or consume those instances. The Service-Finder Ontology was developed based on the requirements we identified by manually inspecting numerous Web sites that publish WSDL as well as many service related information. We made an effort to capture what most of the providers really publish on the Web to convince potential customers to use the services they offer.

The Service Crawler component searches for and retrieves WSDL files, and it also looks for other kind of service-related documentation (service or provider details, pricing information, examples and suggestions from service wikis, etc.); in this way, the crawling operation gathers all relevant data from which deriving the service characteristics important from the users' point of view.

The Automatic Annotator extracts structured information out of the crawled data. It analyzes the output of the SC and creates RDF instances of services, providers and related concepts with respect to the Service-Finder ontology.

The extracted data are then stored and indexed in the Conceptual Indexer and Matcher, which is composed by a knowledge base for the structured descriptions and by a full-text index to allow run-time query of the relevant documentation.

To leverage on the "wisdom of the crowds", the interface of the whole system is offered as a Web 2.0-like site, namely the Service-Finder Portal, which not only provides the usual searching capabilities of a search engine. The SFP allows the system users to proactively contribute to the service description, by adding tags, categories, ratings and comments, free text description and so on to complement the automatically-extracted data.

Finally, a Cluster Engine component collects and analyzes the logs of users' interactions with the SFP to provide service recommendations to the users, letting implicit relations between services emerge from the way people search and browse them.

Results obtained

The Alpha version of the Service-Finder Portal is live since December 2008 and the Beta version is live since October 2009. Even if the project will terminate in 2009, partners aim to keep the portal up and running until 2012. Beta version is accessible at http://demo.service-finder.eu.

In order to evaluate the work done to build the Service-Finder semantic search engine for Web services, we performed several tests divided in two categories: component-level tests and user survey. While the component-level evaluation prove the effectiveness of each part of the system in terms of quantitative and qualitative measures, the user survey demonstrates the interest and the level of satisfaction of users with regards to the whole Service-Finder semantic search engine. More details and further evaluation results are offered in dedicated project deliverables D7.3, D7.5[7].

As result of the project, involved partners aim to exploit individual components both in terms of further research activities but also in terms of consultancies and trainings. Another interesting result of the project is represented by the data collected and produced: information about services, automatic extracted information and service categorization is made publicly available under the *Creative Commons by-nc-sa-unported license*[8].

The Service-Finder project actively co-operated with other European projects and players in the project area. Two of the consortium partners, seekda and the University of Sheffield are members of STI International (http://www.sti2.org/), a collaborative association of interested scientific, industrial and governmental parties that share a common vision in the Semantic Technologies area. Service-Finder co-operates with the NeOn project (http://www.neon-project.org/web-content/) in that in the project the NeOn Toolkit is applied for building the Service-Finder ontologies. The Service-Finder and TAO (http://www.tao-project.eu/) projects met at ISWC 2008 to share results and demonstrate software. TAO representatives presented an overview of that project as well as some of their software, and Service-Finder presented its intermediate results.

Two of the consortium partners, namely seekda and CEFRIEL are also part of the consortium of the SOA4All project (http://www.soa4all.eu/). Service-Finder shares its categorization data with the SOA4All project. Moreover the Service Crawler work within Service-Finder is followed up by the Service location work in SOA4All.

The project targeted and achieved some of the most important dissemination opportunities in the area of Web Services and Semantic Web. This list includes the *International Conference on Service Oriented Computing (ICSOC) 2009* (http://www.icsoc.org/), the *European Semantic Technology Conference (ESTC) 2009 (*http://www.estc2009.com/), the EU Matchmaking Event co-located with the *European Semantic Web Conference (ESWC) 2009* (http://www.sti2.org/events/details/39-the-1st-eu-matchmaking-event-creating-opportunities-for-the-future*), the International Conference on Business Information Systems (BIS) 2009* (http://bis.kie.ae.poznan.pl/13th_bis/)*, a Workshop on Web Personalization, Reputation and Recommender Systems at the International Conference on Web Intelligence (WI) and Intelligent Agent Technology (IAT) 2009 (http://www.wprrs.scitech.qut.edu.au/)* and the *Future Internet Symposium (FIS) 2009:* http://www.fis2009.org/*).* Thanks to these events the innovation aspects of Service-Finder obtained a good visibility.

Further to these international events, the project was present at other national events such as the *Symposium on Advanced Database Systems 2009* (http://www.sebd.org/).

---

[7] All project deliverables are accessible from http://www.service-finder.eu

[8] http://en.wikisource.org/wiki/Creative_Commons_Attribution-NonCommercial-ShareAlike_3.0_Unported

Future Research Directions and Exploitation Prospects

Future works will include other research and development mainly performed outside the scope of the project:

- for the Service Crawler, we aim to further evaluate our focused crawling approaches, especially the new features integrated during the second year of the project.
- for the Automatic annotator, we will make further refinements in preparation for the final release, and use the technology developed in the project to enhance GATE and in other projects, (such as NeOn);
- for the Conceptual Indexer and Matchmaker, we aim to "feedback" research results, like the RDF interfaces we developed in Service-Finder, into the standard OntoBroker;
- for the Service-Finder Interface, we aim to proceed our research efforts in semantic navigation frameworks and web 2.0 frameworks. We are also investigating exploitation prospects in order to integrate our approach and results into existing Web frameworks;
- for the Cluster Engine, we aim to investigate ways to exploit content-based information in order to overcome some of the limitations of collaborative filtering.

Achievements of the Service-Finder project are exploitable in different scenarios:

- Web 2.0 developers may use Service-Finder for building mash-ups of lightweight services.
- ICT Companies engaged in Application Integration scenarios in large enterprises may actively use Service-Finder to look for "address validation", "fraud detection", "credit worthiness", etc.
- The Service-Finder technologies may be sold as an Appliance capable of providing services search capabilities at very low installation and maintenance costs, due to the automated crawling and annotation features.

Further Information

For further information, please visit http://www.service-finder.eu or contact the coordinator of the Service-Finder Project: emanuele.dellavalle@cefriel.it

Project Consortium Information

CEFRIEL SCRL. http://www.cefriel.it/	CEFRIEL	Emanuele Della Valle CEFRIEL SCRL. Milano, Italy E-mail: emanuele.dellavalle@cefriel.it
Seekda OG http://seekda.com/	Seekda	Holger Lausen Seekda OG Innsbruck, Austria E-mail: holger.lausen@seekda.com
Ontoprise GmbH - Intelligente Lösungen für das Wissensmanagement http://www.ontoprise.de/	Ontoprise	Michael Erdmann Ontoprise GmbH Karlsruhe, Germany E-mail: erdmann@ontoprise.de
The University of Sheffield http://www.shef.ac.uk/	USFD	Adam Funk The University of Sheffield Sheffield, United Kingdom E-mail: a.funk@dcs.shef.ac.uk