AN ARCHITECTURE INTEGRATING SEMANTIC E-GOVERNMENT SERVICES

Jan Kolter[1], Rolf Schillinger[1], Wolfgang Dobmeier[1], Günther Pernul[1]

Abstract – Due to the increasing number of citizens using electronic services from home, today more and more federal and local authorities offer their services online. In conjunction with the increasing pressure to save costs, e-Government will continue to grow in the next years. Unfortunately, existing solutions do not provide any semantic information of the offered services, nor a uniform service representation. Pursuing the vision of a complete, European-wide infrastructure, we propose an e-Government architecture that enables service providers to annotate their services with semantic information and that facilitates easy location of relevant e-Government services based on specific life or business events. The presented solution is built on the Service Oriented Architecture paradigm and extends it with Semantic Web and Peer-to-Peer technologies to create a semantic de-centralized service architecture. The proposed architecture was developed within the scope of the EU-funded project Access-eGov.

1. Introduction

The number of e-Government services for citizens and companies is growing constantly. This trend brings advantages for both service providers and service consumers. As the widespread budget deficits of authorities dictate lower payroll costs and streamlined processes, the switch to electronic services is an easy way to automate processes that originally cost a considerable amount of paperwork. Additionally, these electronically offered services give service consumers - like citizens and companies - the freedom to access and use the services from any place at any time. The completion of an authority process turns from an annoying, cumbersome endeavor into an easy-to-handle, time-saving job.

The above-mentioned incentives, however, led to a growing number of e-Government service providers and, subsequently, numerous solutions for an e-Government architecture. Following different standards and frameworks, each authority built an infrastructure tailored just for its individual needs, completely ignoring issues like user-friendliness, interoperability, and service location. The following example demonstrates the deficiencies of existing e-Government architectures:

Suppose, a German resident's driver's license is stolen. Depending on the county or even the city she lives in, she might have the opportunity to report that incident using a web application. Wherever this is not possible, she still could have the opportunity to fax her loss. Maybe though, if there is no other way, she has to file her complaint at a police station in person. In any case, she has to apply for a new driver's license subsequently. Again she will

[1] Department of Information Systems, University of Regensburg, D-93040 Regensburg, Germany
{jan.kolter, rolf.schillinger, wolfgang.dobmeier, guenther.pernul}@wiwi.uni-regensburg.de

examine her options to do this online or in person. Conclusively, finding the correct processes to complete this simple task is not trivial.

The example shows that existing solutions lack a single point of access where citizens are guided through the entire service process. As in most cases a citizen's need involves more than just one e-Government service, a next-generation e-Government solution should also provide an instance, where a citizen's life event is matched to relevant services that are processed in the right order. This vision is the goal of the international project Access-eGov.

Access-eGov is a EU-funded project (2006 - 2008) of the Sixth Framework Programme in the Information Society Technologies activity area. The project consortium consists of 11 partners from 5 different countries, including 3 universities, 3 software development companies, and 5 user partners. The project's main goal is to employ Semantic Web and Peer-to-Peer technologies to build a service oriented e-Government architecture that overcomes the drawbacks of prevailing e-Government architectures and paves the way for a uniform European-wide e-Government infrastructure.

In this paper we introduce a model for a flexible service oriented architecture that overcomes the above-mentioned restrictions of existing solutions and follows the goals of Access-eGov. The remainder of this paper is organized as follows: After collecting requirements in Section 2, we model a reference architecture in Section 3 that uses the imposed requirements to build a flexible, de-centralized service infrastructure. Section 4 presents related work, before we conclude our research in Section 5 and give an outlook on future work.

2. Requirements

A service architecture that implements the goals of Access-eGov must go beyond existing e-Government solutions and should incorporate - apart from well established standards like SOA [1] - new emerging technologies like the Semantic Web and Peer-to-Peer networks [2] [3], which have to be extended and adjusted to build a strong, reliable architecture.

In an effort to design a suitable infrastructure that handles the challenges of Access-eGov, we defined a set of requirements that are listed below:

1. High Availability
2. De-centralized Management
3. Semantic Description of Web Services
4. Composition of Workflows
5. Semantic Mapping
6. Security

The scenario of a semantic, trans-European e-Government platform will require a large-scale architecture that is available at any time. As the platform will include numerous e-Government service providers and even more citizens from several European countries, the architecture should be failure-resistant and not incorporate bottlenecks for services. Moreover, it should allow the flexible allocation of responsibilities. Peer-to-Peer technologies offer the best features to implement these requirements.

The need for high availability goes hand in hand with a de-centralized management. As mentioned above, service providers and consumers may be located in different European countries. In order to guarantee a stable operation of the system, we propose a de-centralized management that facilitates service providers from different domains to manage and maintain the infrastructure independently. De-centralization is also ideally achieved by the application of Peer-to-Peer technologies. This is in contrast to the architecture developed within a similar project we were involved in, the Webocracy project [4], in which client/server technology was used as the underlying architectural paradigm.

The project Access-eGov envisions the idea of a Personal Assistant (PA), which serves as a citizen's interface to the infrastructure. This PA recognizes personal life and business events and matches them to relevant e-Government services. For this purpose, the PA will look up registries, where each semantically annotated e-Government service is listed. This semantic information enables the PA to find and match relevant Web Services to a user's certain life or business event. In order to implement such a scenario, the architecture must provide ways to annotate Web Services as well as an ontology that controls the vocabulary. As most e-Government services are not built from scratch, there must be a way to integrate them ex post, using an instance that is responsible for the wrapping of legacy services.

As mentioned in Section 1, a certain life or business event will most likely require more than one e-Government service. For this reason, the PA should be supported by an ontology that provides the correct compositions of workflows for each event. Following that information, the PA can locate all relevant Web Services and put them the right order.

Semantic annotation and the ability to compose Web Services in an ontology build a very powerful Web Service architecture. However, to make this architecture work in a large-scale environment, the architecture of Access-eGov should be capable of handling the semantic mapping of descriptions from different domains. Service providers from different countries will most likely not use the same terminology for the annotation of their services. To overcome this inconsistency, we propose the use of ontology mediators that perform the mapping between certain domains.

As the interaction with public authorities include sensitive data, special care must be taken while addressing security issues like identification and authentication services. Additionally, access control and auditing facilities must be provided as well as non-repudiation and anonymity.

In recent years, an increasing number of discussions about the use of open standards helped raise public awareness about this issue. On behalf of the EU, the Danish government defined the term open standard in [5]. According to this definition a standard qualifies as an open standard, if it is accessible to anyone free of charge and includes detailed documentations about all its functions. Furthermore, in an effort to support decision makers in the field of software architectures, the EU has set up an online "reference profile"[2], which lists more than 600 relevant standards, each with a short description and a status attribute. Even though this source is not limited to open standards, it is still a valuable resource to find applicable standards for implementation decisions.

Considering all requirements, it is evident that a derived architecture will exceed state-of-the-art solutions and will have to include modern developments mentioned above. The following

[2] http://standarder.oio.dk/English/

section will present a derived architecture, which meets all requirements defined in this section.

3. Proposed Architecture

Developing a solution for our scenario, we modeled an architecture depicted in Figure 1. A main pillar of our proposed work is a Peer-to-Peer network. Such networks gained considerable popularity in recent years due to the following developments: First, Peer-to-Peer based file sharing networks raised the attention of the media, by quickly turning into huge illegal distribution networks of copyrighted content. On the other hand, a lot of scientific effort has been put into the research of this technology.

As according to [3] Peer-to-Peer networks are "capable of adapting to failures [...] without requiring the intermediation or support of a global centralized server or authority", they exactly fulfill the requirements High Availability and De-centralized Management imposed in Section 2. Common implementations of this technology are Jxta[3], P-grid[4] and Oceanstore[5].

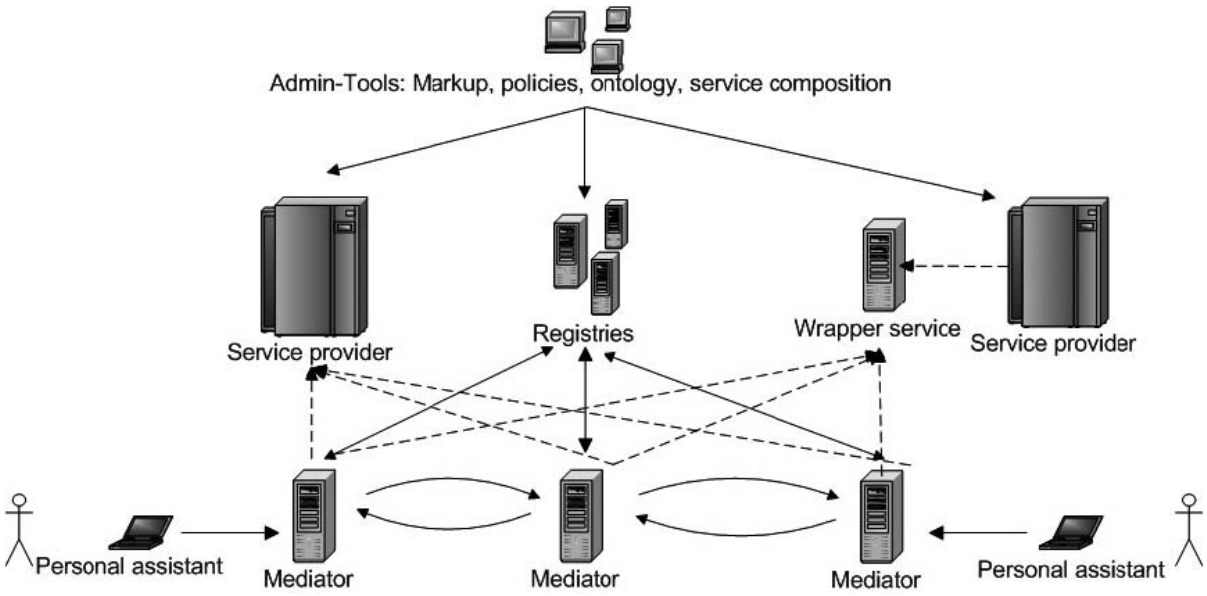

Figure 1: Architecture for Integrating E-Government Services

Another important constituent of our proposed architecture is the Personal Assistant (PA), which can be implemented on many different platforms, ranging from mobile devices like PDAs and cellular phones to desktop PCs. The PA's main task is to query the infrastructure and present the corresponding results to the user. Performance constraints, however, disallow making each PA part of the Peer-to-Peer network. Instead, we use an extended form of a well-known design pattern: the mediator [1]. In our proposed architecture we consider mediators to be responsible for the communication between participating nodes. The motivation for this solution was the prevention of the otherwise intense growth of Peer-to-Peer traffic. In general, mediators are responsible for the following tasks:

---

[3] http://jxta.org
[4] http://p-grid.org
[5] http://oceanstore.cs.berkeley.edu

- **Interface to Service Providers and Legacy Services**

  The source of all information in the infrastructure is the service provider, which represents either a newly designed and thus already semantically annotated Web Service, or an already existing legacy application. The defining attribute in this context is semantic annotation. To open the Access-eGov architecture for legacy systems, we need to access them by means of a customized wrapper. Such a wrapper must provide two separate layers, one that interacts with the infrastructure, and the other one that communicates with the service to be wrapped. While the first layer is identical for all wrappers, the second layer is tailored to fit exactly the needs of the already existing legacy system. Semantic annotation in this case is done via external applications, e.g. a web interface provided by the wrapper.

- **Management of Life and Business Events**

  The main objective in the architecture is carrying out a user task. In some cases, there might be a single service whose goals exactly match the user task. Another possibility is that one service is not sufficient for fulfilling the task and a process chain has to be executed instead. Process chains that represent workflows can consist of online and offline services. The automatic discovery of such a workflow is possible as long as the user task is not too complex and all involved services are online. Process chains that either contain offline services or services exceeding a certain complexity will always have to be manually composed.

- **Management of the Registry Infrastructure**

  The use of Peer-to-Peer networks usually comes with a large performance loss. For this reasons a careful planning of the network infrastructure is indispensable. As the most frequent operations in this network are service queries, we have to adjust our network carefully to find relevant registries with as little unsuccessful queries as possible. There are a number of ways to store content in Peer-to-Peer networks. In [3] those opportunities are evaluated but the results can not be accurately applied to our specific scenario of storing and quickly locating relevant references to web services. In order to facilitate the efficient service storage and location, we propose the partition of service descriptions according to certain criteria. This will contribute to a balanced network load between participating nodes.

- **Interface to Personal Assistants**

  One of the most important decisions in designing this architecture is the allocation of responsibilities between mediators and PAs. Slow and featureless thin clients, like those running on PDAs or cellular phones, have low computing power and usually slow network links, and thus demand that most of the processing has to take place in the mediators. On the other hand, a full-blown client application on a desktop PC can handle much of that processing. Providing valuable details about the users' client devices, user requirement analyses within the scope of the project Access-eGov will help us find the right balance of responsibilities.

It is evident that mediators are the most complex parts of our architecture. For this reason, we recommend implementing those mediators as closely coupled modules.

In order to guarantee a stable flow of information, participating Web Services have to be semantically annotated with a common vocabulary provided by a specific ontology. The

ontology of our choice is WSMO[6], which - unlike OWL-S[7] - was designed from scratch addressing exclusively the needs of semantic web services. WSMO inherits four core elements from the Web Service Modeling Framework [6], an earlier research project. According to [7], the main components of WSMO are Ontologies, Goals, Web Services and Mediators. An extensive comparison of WSMO and OWL-S is beyond the scope of this document. The interested reader may refer to [8].

As the data transferred over this infrastructure might be of very private nature, powerful security facilities have to be put in place. All transport channels need to be encrypted and access control should be configurable on a per-service or per-process-chain basis. The underlying authorization schema will be a variant of the Attribute-Based Access Control (ABAC) [9], which specifies authorizations of subjects on objects based on certain attributes, and not necessarily on the identity of requesting users. This mechanism especially simplifies the administration of privileges in an e-Government environment, as not all subjects may be known to the Security Officer in advance. The application of the traditional Role-based Access Control would result in an enormous number of different roles [10]. Attributes for ABAC could be supplied in form of Identity and/or Attribute Certificates.

Furthermore, special attention has to be put on the identification of users. The lack of availability of an EU-wide smartcard in this context is a serious obstacle on the way to a user-friendly solution. Another promising approach is the integration of user certificates into browsers, which is technically possible but currently far from being user-friendly.

Communication security and the authorization of users are only parts of the overall system security. For example non-repudiation of certain communications demands comprehensive logging facilities in all involved components. Additionally, extensive auditing facilities keep up a maximum security level by reporting configurable incidents to selected administrators.

In order to reuse as much existing code as possible, we carefully investigated available technologies, putting strong emphasis on their openness. Thanks to the ever growing open-source community, there is a large repository of reusable modules for all kinds of purposes. With the help of the already mentioned EU reference profile we envision the use of the following technologies for the implementation of our architecture:

- **Communication**
  The Internet Protocol (IP) serves as the basis our chosen Peer-to-Peer framework operates on. JXTA or a modified variant of Java Spaces[8] are the most likely candidates for a Peer-to-Peer framework.
- **Semantic Annotation**

  Earlier in this article, we already have laid out our reasons for choosing WSMO as our ontology. Therefore, using WSML[9] as complementing markup language is a logical choice. Presumably, during the course of implementing our architecture, WSMX[10] will be extended to fit our exact needs. The WSM family in turn is based on the well-known web service paradigm, which is approved by the reference profile.

---

[6] http://wsmo.org
[7] http://www.daml.org/services/owl-s/1.0/
[8] http://java.sun.com/docs/books/jini/javaspaces/
[9] http://wsmo.org/wsml/
[10] http://wsmx.org

- **Security**
  Securing communication channels with either IPSEC or SSL is already a common practice. In this context, X.509 standard is a de-facto standard. The reference profile does not mention any alternatives to X.509, as it only lists standards that facilitate the exchange of security information and the RBAC model.

4. Related Work

Various proposals of service-oriented infrastructures for e-Government have been made. Gouscos et al. [11], to the best of our knowledge, were among the first to introduce a one-stop architecture employing mediators and a workflow engine for the composition of services. However, they do not employ semantics for the location of services nor Peer-to-Peer concepts for the ad-hoc extension of the infrastructure.

Contenti et al. [12] present a distributed architecture for service orchestration. They also pursue the integration of legacy services. In so doing, they define several complex components (Cooperative Gateway, Information Manager and Presentation Layer) that participating authorities have to deploy on their sites in an effort to build a uniform environment. In our approach, authorities only have to use a simple wrapper service, which can even reside on a remote location. No modifications to the existing infrastructure are necessary.

The Qualeg Project[11] is a project dealing with the issue of Quality of Service in e-Government. In this project, facilities for assessing citizens' opinions about offered services were developed. This includes ontology-driven workflow and semantic engines for the composition and publishing of annotated services.

Apostolou et al. [13] propose a three-layer logical architecture, consisting of modeling, configuration and runtime components. The goal of this approach is to manage the life cycle of e-Government services, i.e. the service design and configuration.

Finally, [14] is a proposal similar to the one presented in this paper. However, they do not use distributed service registries and Peer-to-Peer networking.

5. Introduction

The widespread use and acceptance of e-Government services will strongly rely on a common, fault-tolerant architecture that empowers authorities to merge their services into a uniform e-Government infrastructure. Moreover, easy access and user-friendliness for citizens and companies are decisive success factors.

In this paper, we present a novel innovative e-Government architecture, which addresses exactly the above-mentioned prerequisites. Our solution extends the traditional paradigm of the Service Oriented Architecture and uses Peer-to-Peer technologies to improve availability and management in a large-scale environment. Furthermore, we employ Semantic Web technologies for the annotation, location, and composition of e-Government services. We also consider the semantic mapping of services from different domains. Finally, we discuss

[11] http://www.qualeg.eupm.net

security requirements to our proposed architecture. Future work will include the creation of a reference implementation within the scope of the project Access-eGov. This effort will also involve the careful testing in selected pilot scenarios.

Acknowledgment

We would like to thank our project partners for helpful comments and stimulating discussions. This work is done within the Access-eGov project, which is supported in part by the European Union under the IST Programme, contract No. FP6-2004-27020. The content of this publication is the sole responsibility of the authors, and in no way represents the view of the European Commission or its services.

References

[1] Gamma, E., Helm, R., Hohnson, R., Vlissides, J.: Entwurfsmuster. (Addison-Wesley 2001)

[2] Lee, T.B., Hendle, J., Lassila, O.: The Semantic Web. Scientific American (2001)

[3] Androutsellis-Theotokis, S., Spinellis, D.: A Survey of Peer-To-Peer Content Distribution Technologies. ACM Computing Surveys 36(4) (2004) 335–371

[4] Paralic, J., Sabol, T., Mach, M.: A System To Support E-Democracy. In: Proc. of the 1st International Conference on Electronic Government (EGOV 2002), Aix-en-Provence, France (2002)

[5] The Danish Ministry of Science Technology and Innovation: Definition of Open Standards (2004)

[6] Fensel, D., Bussler, C.: The Web Service Modeling Framework WSMF. Electronic Commerce Research and Applications 1(2) (2002) 113–137

[7] Feier, C., Domingue, J.: D3.1 v0.1 WSMO Primer. (2005)

[8] Lara, R., Polleres, A., Lausen, H., Roman, D., de Bruijn, J., Fensel, D.: A Conceptual Comparison Between WSMO and OWL-S. (2005)

[9] Priebe, T., Fernandez, E., Mehlau, J., Pernul, G.: A Pattern System For Access Control. In: Proc. of the 18th Annual IFIP WG 11.3 Working Conference on Data and Application Security, Sitges, Spain (2004)

[10] Priebe, T., Dobmeier, W., Muschall, B., Pernul, G.: Ein Referenzmodell für attributbasierte Zugriffskontrolle. In: Proc. of the 2. Jahrestagung Fachbereich Sicherheit der Gesellschaft für Informatik (Sicherheit 2005), Regensburg, Germany (2005)

[11] Gouscos, D., Laskaridis, G., Lioulias, D., Mentzas, G., Georgiadis, P.: An Approach To Offering One-Stop E-Government Services - Available Technologies and Architectural Issues. In: Proc. Electronic Government: First International Conference (EGOV 2002), Aix-en-Provence, France (2002)

[12] Contenti, M., Mecella, M., Termini, A., Baldoni, R.: A Distributed Architecture For Supporting E-Government Cooperative Processes. In: Proc. E-Government: Towards Electronic Democracy, International Conference (TCGOV 2005), Bolzano, Italy (2005)

[13] Apostolou, D., Stojanovic, L., Lobo, T., Thoenssen, B.: Towards A Semantically Driven Software Engineering Environment For Egovernment. In: Proc. E-Government: Towards Electronic Democracy, International Conference (TCGOV 2005), Bolzano, Italy (2005)

[14] Sabucedo, L., Rifón, L.: A Proposal For A Semantic-driven Egovernment Service Architecture. In: Proc. Electronic Government: 4th International Conference (EGOV 2005), Copenhagen, Denmark (2005)