Applying TOGAF to Define and Govern a Service-oriented Architecture in a Large-scale Research Project

Aneta Kabzeva¹, Michael Niemann², Paul Müller¹, Ralf Steinmetz²

¹University Kaiserslautern
²TU Darmstadt, Germany
{kabzeva; pmueller}@informatik.uni-kl.de
{michael.niemann; ralf.steinmetz}@kom.tu-darmstadt.de

ABSTRACT

The Service-oriented Architecture paradigm has become a widely adopted solution for enterprise application landscapes. SOA promises system integration improvement, business and IT alignment, applications reusability, and fast adaptability to changing requirements. Yet, the advantages that SOA offers lead to its main management challenges. An SOA consists not only of services. Business requirements, processes, models, rules, policies, and different stakeholders play also an important role. To manage this increased complexity and heterogeneity in SOA landscapes enterprises apply architecture frameworks and governance approaches. An essential enabler for SOA Governance is the traceable documentation of the service-oriented application landscape. This paper shows how the TOGAF Architecture Development Method was applied in a large-scale SOA-based research project to support the governance proposes in the context of a service marketplace. Supplementary templates were defined to guide the description of services, data models, and architectural components. Some of the models and decisions proposed by TOGAF were shifted to other phases to fit the innovative nature of research projects.

Keywords

Service-oriented Architecture, Governance, Architecture, Research project, TOGAF.

INTRODUCTION AND MOTIVATION

Service-oriented architecture (SOA) has set up in the past years as a concept for the development of adaptive software landscapes. Service-oriented systems are developed by combining atomic, loosely coupled and distributed software components, referred to as services (Erl 2005). The flexible composition of services shared across organizational boundaries advocates the reusability of applications and boosts business agility. Promising a better alignment of business and IT, the SOA paradigm has become an important concept for structuring enterprise architectures. The adoption of SOA allows a better system integration and reaction to changing business requirements but it introduces also new challenges. New roles, new development task, and new artifacts have to be considered. To manage this complexity a traceable documentation of all service-oriented application landscape artifacts is needed.

Large heterogeneous projects with numerous partners give rise to a high complexity, which is comparable with the complexity of enterprise architectures. Compared to the development of an enterprise architecture in a large-scale research project we consider the separate parts of all the different project partners and their interactions rather than the entirety of one enterprise. However, there are similar inherent structures in both, companies and well-organized large research projects. The architecture management within a large-scale project with numerous legally independent project partners means harmonization of different processes, architectural views, stakeholders, data models, technologies, application landscapes, and systems to work as a whole for the realization of an innovative vision. In enterprise environments different architecture frameworks, i.e. Zachman (Zachman 1987), TOGAF (The Open Group 2009a), DoDAF (DoD 2009), are applied to control the landscape complexity. In his last version TOGAF even offers a SOA extension. For complex research projects this is not yet the case. We argue that the methods provided by enterprise architecture development can be applied within large heterogeneous research projects to structure the architecture design and management process, improve the traceability of relationships between the different artifacts designed and developed within the project, and last but not least to enhance the communication between the project partners and their understandability on the project. To prove this assertion we have applied The Open Group Architecture Framework TOGAF (The Open Group 2009a) for the architecture design and
Governance in a large research project based on SOA. Having to consider the coordination of a large number of financially and legally independent project partners; the predefined project proposal, goals, and structure; the unexplored influence of novel technologies and methods on the architectural quality; and the development of a prototype, but not a ready to use system as a result, the order of the TOGAF phases and the models implemented within each of them had to be changed a bit. Also, some additional templates have been defined to improve the relationships between the entities developed within the project.

The remainder of the paper is structured as follows. A short overview on the Theseus/TEXO research project on which TOGAF was applied is given in the next chapter. A short discussion on governance challenges in service-oriented landscapes and the place of architectural governance in the governance lifecycle follows. The rest of the paper outlines the different phases of the TOGAF Architecture Development Method, their application order and output within the scope of Theseus/TEXO. Finally we draw a number of conclusions.

**THESEUS/TEXO - BUSINESS WEBS IN THE INTERNET OF SERVICES**

The research program Theseus (Theseus 2009) is funded by the German Federal Ministry of Economy and Technology. Its goal is the development of a new Internet-based infrastructure for better usage of the knowledge available Internet. The program integrates six use cases. The goal of TEXO as part of Theseus is to provide a SOA-centric platform which allows services to be traded and composed in value-added services via the Internet. TEXO involves more than 60 researchers from 15 partners from the industry and academic sector and focuses on web based Business Services (i.e., e-Services) and their automation using Technical Services (e.g., Web Services). TEXO addresses the full lifecycle of these services via intuitive interfaces and technical systems. The lifecycle addresses Service Providers, Hosters and entrepreneurs, offering new services to consumers (individual users as well as organizations) – while stakeholders gain the opportunity to design, compose and provide new services. Meeting all these challenges demands a good managed interdisciplinary approach in order to create a platform supporting all stakeholders and phases of the service lifecycle.

The TEXO team is organized in twelve work packages. Every work package comprises researchers from one or more of the partner organizations according to their domain of interests within TEXO. One of the work packages is responsible for the management of the project. However, all packages work in parallel for the execution of the research activities planned within the project proposal.

**GOVERNANCE CHALLENGES IN SERVICE-ORIENTED APPLICATION LANDSCAPES**

Concerning Governance for Service-oriented Architectures, there are a number of definitions, diverging in focus. The most known definition for IT Governance by Weill/Ross “IT Governance: specifying the decision rights and accountability framework to encourage desirable behavior in the use of IT” (Weill and Ross, 2004) emphasizes the thoughtful assignment of responsibilities and roles in an EA (or SOA) in order to stimulate desirable behavior. Bloomberg defines SOA Governance as the application of IT Governance to an SOA system: “it is how IT Governance should operate within an organization that has adopted SOA as their primary approach to EA.” (Bloomberg, 2004). Niemann et al (2008) define the term “SOA Governance” as follows:

> “SOA Governance is a management model, providing the abilities to guarantee sufficient adaptability and integrity of the SOA system as well as to check services concerning capability, security aspects and strategic business alignment. Its overall goal is SOA Compliance, i.e. compliance to legal, technical and intra-company regulations, respectively. Particularly, it ensures the reliable long-term operation of an SOA.”

According to them, a governance approach aims at smooth adoption and successful operation of an SOA. Main tools are the provision of guidelines and mechanisms to ensure the integrity of an SOA and its adaptability to company-inherent processes. Monitoring and control of services concerning security issues and the alignment to business processes are part of governance tooling. A best practice catalog serves as repository of recommendations. Supporting the achievement of IT goals and realizing business-IT alignment, as well as achieving adherence to regulations and standards (such as, e.g., the Sarbanes Oxley Act) are main goals of SOA Governance.

Referring to the Deming cycle (plan, do, check, act) [Shew39], we distinguish two major phases in service governance: design-time governance (plan and act) and runtime governance (do and check). Afshar defines eight key leverage points for policies: people, financial, portfolios, operational, architecture, information, technology, and project execution [Afsh07]. According to our understanding of design time and runtime governance, we group these leverage points as follows. Runtime governance focuses on operational, project execution, and financial aspects. Design time governance comprises the domains
architecture, information, technology, people, and portfolios. With respect to monitoring (check), parts of information, technology, and architecture are also concerned in runtime governance. Figure 1 illustrates the assignments.

![Design time governance vs. Runtime governance](image)

**Figure 1: Design-time and runtime governance**

**APPLICATION OF THE TOGAF ARCHITECTURE DEVELOPMENT METHOD TO A RESEARCH ENVIRONMENT**

One of the important aspects to be governed during design time of service-oriented architectures is the actual architecture. In the following, we investigate the domain architecture in particular, based on experiences made during the work in the research project Theseus/TEXO. The architecture design and modeling process applied within TEO uses the Architecture Development Method of The Open Group Architecture Framework (TOGAF) framework as basis. Considering the specific characteristics of a research project – heterogeneous financially and legally independent project partners, predefined project structure and work assignments, and the development of innovative technologies - and the collected experience with them, the next chapter presents how the different models can be implemented and at which phase.

The Open Group Architecture Framework (TOGAF) is one of the most popular frameworks for enterprise architecture management worldwide (Keller 2009, Schmelzer 2009). The core of TOGAF is the Architecture Development Method (ADM). ADM is a generic method, which can be applied on a broad range of different architectures (The Open Group 2009a). However, the last version 9 of the framework, which was released in February 2009, gives guidelines on the adaptation of the TOGAF ADM on service-oriented architectures. Regarding the large number of project participants and the complex heterogeneous environment within the Theseus/TEXO project we consider the application of the TOGAF service-oriented ADM appropriate for a large scale research project. In the following we present the application procedure of TOGAF to the situation in the research project. Basically, we compare our procedure with the TOGAF procedure. Additional outputs generated within TEO are listed in the appropriate phases.

**Preliminary Phase**

This first phase of the ADM cycle is there to prepare the enterprise organization, or in our case the project team for the development process. This includes in the case of a service-oriented solution the presentation of the service orientation principles, the consideration of existing reference architectures and models for SOA, and the elaboration of a governance strategy for the planned SOA solution (The Open Group 2009b).

Since Theseus/TEXO is a scientific project, the tasks prescribed by the TOGAF preliminary phase to be used during the research were done during the project proposal definition from every partner involved in the project. This proposal immediately also defined the structure of the project –the assignment of the partial solutions to the most appropriate partners.
according to their domain knowledge. Thus, “the delivery vehicles” identification (The Open Group 2009b) of Phase E: Opportunities and Solutions was shifted to the preliminary phase. Since the main goal of a research project is the definition of innovative combinations of existing standards or the definition of new ones, the management and the implementation plan (Phase F: Migration Planning) were left to the separate project partners, who are also the best specialists in their domain. Given that innovation is a time consuming and laborious process, the realization opportunities for the research ideas run parallel to the architecture definition process. On the one hand they were influenced by architectural requirements and constraints defined during the architecture development. On the other hand the results of the research work, which often lead to the discovery of new requirements and constraints, reflected in the architectural models and views. To summarize, the Preliminary Phase of the ADM in a research environment ends up with more output than originally defined in TOGAF.

Phase A: Architecture Vision

The steps in this first phase of the ADM cycle comprise the definition of the business goals, requirements and constraints. A high-level description of the final architecture has to be presented that shows what will be the business benefits from the future solution. Depending on the identified stakeholders for the system, different architectural views have to be planned for development (The Open Group 2009b).

The definition of the requirements and constraints for the marketplace developed within Theseus/TEXO was executed with the help of scenarios (Kabzeva et. al. 2009). An initial unstructured and informal version of the requirements was already available in the project proposal. Based on the project proposal, every one of the project partners defined his vision on the TEXO marketplace in the form of scenarios (Alexander and Maiden 2004). After analysis of the collected scenarios all requirements placed on the system were collected and discussed. Subsequently, an integrated scenario presenting the set of requirements and constraints that all partners agreed on was elaborated. Since in the unexplored context of a services marketplace did not allow an explicit identification of all system stakeholders, business processes, and data needed a detailed planning on the architectural views to be modeled was also impossible at that stage in a research environment. Three Task Forces were set up to research on these topics – the Methods Task Force (MeTF) in charge of the business architecture, the Architecture Task Force (ATF) in charge of the architectural models, and the Modeling Task Force (MoTF) responsible for the definition of the data structures. Every domain considered within the project was presented by one domain member within every one of the Task Forces.

Additionally to the outputs prescribed by TOGAF for this phase, a project vocabulary (Texopedia) was set up and a TEXO service lifecycle in the context of the service marketplace was defined. The integrated scenario, the TEXO service lifecycle, and the vocabulary were set as guidelines for the Task Forces.

Phase B: Business Architecture

According to the TOGAF specification for SOA, this phase defines five SOA specific business models considering the requirements identified in the Architecture Vision – Business Process Model, Business Roles Catalog, Business Vocabulary, Business Rules Catalog, and Business Services Catalog (The Open Group 2009b).

A Business Process Model has to define the set of business processes relevant for the future system and a three or four levels of decomposition of these processes. In Theseus TEXO, this activity was performed by the MeTF. Considering the integrated scenario and the service lifecycle from Phase A, for each phase of the lifecycle a high-level business process has been defined. Each process has been defined in three levels of granularity or abstraction, tagged with responsible or contact persons. In contrast to the top-down approach pursued in TOGAF, in a project environment no complete definition of the business processes in Phase B was possible. Since all three Task Forces worked in parallel (Figure 2) in the first iteration only an initial business process model was defined. This was later consolidated with the results of the ATF and MoTF to create the Service Architecture in Phase D.

The Business Roles Catalog in TEXO represents all users of the TEXO marketplace with their responsibilities in the business processes, as TOGAF prescribes. However, just like the Business Process Model, this catalog is an initial version in the first iteration, which is to be consolidated with the stakeholder roles defined by the ATF.

A special Business Vocabulary has not been defined in TEXO. However, there is a common vocabulary used in Service-oriented Computing. New terms as well as our understanding of “service” were added to the Texopedia.

The project partners are from different legally independent organizations. The work done from them within the project has to follow the organizational governance rules of the own organization. Of course, at interaction points a common solution has to be found that complies with the rules of all partners involved in the interaction. Such a solution is specific for the case, so no central rules were defined. As for the Business Service Catalog, the innovative nature of a research project, and the fact that
not a ready to use system but a prototype is the result of research projects, no facts on the concrete provider, consumer, or value of the business services could be specified.

![ADM as applied in Theseus/TEXO](image)

**Figure 2: ADM as applied in Theseus/TEXO**

**Phase C: Information Systems Architectures**

According to the TOGAF ADM Phase C defines data sources and types necessary to support the business processes, and the structure of the applications responsible for the implementation of the processes defined in Phase B. As mentioned above, the MoTF and ATF that are in charge of these steps work in parallel with the MeTF. So the Data and Applications Architecture in TEXO were defined only on the basis of the integrated scenario and service life cycle from Phase A.

The Open Group does not prescribe any SOA specific models for the data part of this phase. For the application part eight models are mentioned – Service Interaction Model, Business Process/Service Matrix, Service Consumers Matrix, Service Contract and Policy Catalog, Service Access Control Model, Service Configurations and Provisioning Model, Service Loading Model, and Service/Application Matrix (The Open Group 2009b).

One of the most important aspects in a service-oriented landscape is the service description, which defines how a service interacts with its environment – its functional and non-functional properties. A semantic service description model has been defined by the MoTF, i.e., for the semantic description and retrieval of services, a holistic service description ontology, based on the foundational ontology DOLCE, has been developed (Oberle et. al. 2009). In the context of a service marketplace, another important issue is the consideration of legal aspects for service consumption. A generic service contract template and a model defining laws, which can be included in service contracts, were also created within the MoTF. These template and model substitute the Service Contract and Policy Catalog in TEXO. In a research environment, no service specific contracts and policies can be completely defined before starting with the implementation. Additionally a model description template has been defined to ease the traceability between models, components, and stakeholders.

The TEXO Service Interaction Model that shows how the individual services interact with each other was defined with a meet-in-the-middle approach (Kabzeva 2009, Twidale and Floyd 2008). From the top-down direction ATF defined an architecture based on the integrated scenario and the service lifecycle from Phase A. From the bottom-up direction, regarding the package based structure of the project; every package defined a package specific architecture. The resulting ten
architectural views were consolidated after analysis. The consolidated Service Interaction architecture comprised all services planned for implementation from all project partners. Regarding the functionalities these services offer and the possible interaction points with the marketplace, according to the architecture, a generic stakeholder model was defined. Another valuable output from the work within the ATF was a component description template to ease the traceability between components, processes, and architectures.

The Business Process/Service Matrix was considered after the consolidation of the results from phases B and C. No separate Service Consumers Matrix was defined. The roles were inserted on the respective places in the architectural view since a “picture” is more intuitive that a matrix with more than 200 entries. Service Access Control Model, Service Configurations and Provisioning Model, and Service Loading Model were not defined until that point of time. Access control will be inherent to service development and rather in the scope of the service developing role. Together with service configuration and provisioning, these two issues are elaborated on as soon as the marketplace is set up. A Service/Applications Matrix is needed only in case existing applications are wrapped as services, which was not the case in TEXO.

Phase D: Technology Architecture

The purpose of this ADM phase is the definition of the infrastructure for the SOA solution, i.e., the infrastructure needed for support and execution of the defined services. The following three models have to be defined – Technology Portfolio, Service/Physical Systems Matrix, and Service/Technology Matrix (The Open Group 2009b).

Within TEXO Phase D is referred as Service Architecture because it consolidates the results from Phase B and C before defining the Technology Architecture. Via the definition of a Business Process/Service Matrix a gap analysis is performed between the processes from Phase B and the services from the Service Interaction architecture of Phase C. Thus missing services or (parts of) business processes were identified and the respective process and architecture models were consequently updated. The generic stakeholders and Business Roles models were also consolidated in one model with two abstraction levels. Considering the consolidated TEXO Roles Model for every main role a role-specific architecture view was defined by the ATF. A role-specific view shows only the architectural components needed for the concrete role to perform its role-specific functionalities on the TEXO platform. For every complex process initiated by a role interaction a process-specific view showing how the architectural components interact with each other to perform the process were also modeled.

To improve the visibility within the architectural views and to show the assignment of the numerous services to the physical systems in a low-complexity way, a color coding system was set up. Every architectural component was dyed in a specific color depending on the physical runtime it has to be deployed on and the functionality it plays. Thus no explicit component’s type notification in the architecture is needed without loss of information. Our color coding system replaces the Service/Physical Systems Matrix proposed by The Open Group. A Technology Portfolio and the following Service/Technology Matrix do not exist until now. The completion of these two is still in progress, as the project is still running and the researchers are working on the technologies for their components.

Phase E: Opportunities and Solutions

The Opportunities and Solutions phase is there for the identification of the delivery mechanisms for the architecture planned in previous phases of the ADM. Particularly in the context of SOA this means the identification of projects within which the services will be implemented and the decision which parts of the target architecture will be implemented in-house and which will be outsourced (The Open Group 2009b). In the context of a research project this corresponds to the decision on which parts of the planned architecture will be implemented by which project partner. This decision is taken already with the definition of the project proposal, which is part of the Preliminary phase in that case. The decision if the separate project partners will implement their partial solution on their own or will use external or existing services is left to themselves. Thus no extra Opportunities and Solutions phase is needed in a research project.

Phase F: Migration Planning

The output of Phase F is a detailed plan for the implementation of the architecture. In a research environment the implementation of the actual services is left to the responsible project partners. The only constraint is that the results have to be compliant with the requirements set in the project proposal and arranged with all project partners. This means as a consequence, the governance model that has to be reviewed within this phase (The Open Group 2009a), has to be set as a central issue in a research project. Since the implementation plan is set by the separate project partners alone in compliance with both the projects and their own organizational governance rules, no extra Migration Planning phase was introduced within TEXO.
Phase G: Implementation Governance

Ensuring that the system is implemented as planned in the architecture is important for a service-oriented landscape as for any other system. Given the fact, that in a research project the requirements and the architecture mature with the implementation process to the same extent as they govern it, the governance in the TEXO ADM was shifted to the center. On the one side the governance in a research project sets initial constraint that have to be followed from the beginning of the project, because of the specific structure of a research project. On the other side the realization process helps the governance process to grow, to detect new governance areas, and to improve its policies descriptions. Even though already recognized as central aspect, the governance framework set up within TEXO (Niamann et al. 2008) is still work in progress that will be extended in the forthcoming integration phase of the partial solutions.

Phase H: Architecture Change Management

Consideration of change recommendations for the architecture and the architectural process which is the purpose of phase H, happen in a research project during the whole lifecycle of the project. However, the changes are not done as a response to the performance assessment of the solution, but to the research in progress results. Design and development walk hand in hand to realize the most appropriate solution for the problem described in the project proposal and thus learn continuously from each other. In a research environment Architecture Change Management comprises not only the improvement steps after the implementation of the target system, but is a central aspect, which improves continuously the quality of the project with the growth of the research results. The main goal of a research project is the development of innovative combinations of existing standards and technologies or completely new of that ilk. During the innovation process new possibilities or constraints on the researched system can be recognized. Thus, the (low-level) processes and the architectural views have to be dynamically adapted to the changing requirements. Change management with respect to performance assessment of the TEXO prototype is future work.

CONCLUSION

In this paper we showed how we applied TOGAF to manage the complexity within the large-scale project Theseus/TEXO. We described the experience we made with the ADM cycle and models prescribed by The Open Group for a service-oriented environment and how they were realized within the project. The Preliminary Phase in a research environment generates more output than in an enterprise context, since a project proposal already defines the structure of the work and the implementation responsibilities of the partners involved. Thus, decisions taken normally in Phase E and F are shifted forward in the cycle. The parallel work of all project partners within the different domains did not allow the specification of the business processes and the corresponding architecture before starting with the implementation works. On the one side this can be captured as negative, regarding the updates needed by interaction or conformance problems. However, regarding the innovative nature of research work, parallel design and development is inevitable. Thus, Phase A defined additionally to the architecture vision three task forces to govern and consolidate the parallel work process and the outputs produced in the not subsequent as prescribed by TOGAF, but parallel Phases B and C. As a regulator for the work of the three Task Forces Phase A has to deliver a project-specific service lifecycle as output. The consolidation work is done in regular intervals as part of Phase D. To support the parallel work in a research project and its innovative nature, the governance and change management phases have to be shifted to the center of ADM.

In addition to The Open Group output prescriptions from the separate phases, we considered the set up of three Task Forces to serve as governance boards for the parallel work within the project and look after the development and adaptation of the business, data, and application artifacts. A project-specific service lifecycle has to be set up in addition to the business requirements on the desired result. Rather than a top-down approach, a meet-in-the middle architecture design approach is the best alternative in a research environment. This guarantees the consideration of all work package and domain-specific views on the future solution. Since a research project develops a prototype and no ready-to-use system, no explicit definition of the service owners, users, and their contracts is possible or expected. More useful in this case are templates defining the possible structures for these documents. Supplementary templates for the description of services, law aspects, data models, and architectural components proved a valuable output within TEXO.

DISCLAIMER

The project was funded by means of the German Federal Ministry of Economy and Technology under the promotional reference 01M07012. The authors take the responsibility for the contents.
REFERENCES

12. The Open Group (2009a) TOGAF Version 9, Van Haren Publishing