Chaining Environmental Web Services: Composing a Web Service Chain

HUUB SCHOLTEN
Wageningen University, The Netherlands

ADRIE J.M. BEULENS
Wageningen University, The Netherlands

AYALEW KASSAHUN
Wageningen University, The Netherlands

ARNO KRAUSE
Alterra, The Netherlands

Keywords: chaining of Web services, model-based decision support, environmental application domain, semantic interoperability, ontology, defining chain functionality

Chaining environmental Web services

Decision making in the context of environmental problems is extremely complex, as the multifaceted interdependencies between relevant physical and man-induced processes need to be considered simultaneously. Interactions between these processes make the problems an all encompassing threat, as realised by the global climatic change, to which policymakers and scientists must join forces.

To support solving environmental problems and problems from other application domains with a similar complexity, all kind of modules are used to build decision supporting systems. These loosely or tightly coupled components can include one or more models, data and tools for pre-processing data, visualisation, statistical analysis. Building monolithic or client-server architecture decision supporting systems for a specific problem (problem type, location and time frame) requires substantial efforts, including collecting the modules, technical coupling, organising proprietary issues and arranging the proper expertise to use them.

Nowadays, many organisations that are involved in solving complex environmental issues (or similar other ones) consider to move from selling (modules for) such decision supporting systems to providing their functionality as service instead. There are several reasons to do so. First the ease of maintenance, as all services are unique and no copies are needed for specific decision support systems. Furthermore, these organisations can use these services ubiquitous for their own purposes and sell services to others. In this way the organisations can easily manage their proprietary tools and data. The final argument to migrate to a Web service approach is to facilitate interoperability between components that run on different platforms.

Several initiatives aim at better using and sharing information, data and tools. Achieving a Single Information Space in Europe for the Environment (SISE) is supposed to be a major initiative to realise such an information infrastructure. Other initiatives, e.g. the Shared Environmental Information System (SEIS) and its backbone, the Infrastructure for Spatial Information in Europe (INSPIRE) are main contributions to the infrastructure aimed by SISE, but these initiatives mainly focus on data collection and data sharing and data on itself does not lead to solutions for the environmental problems or to adaptation to climate change. Models, applications and services are also needed. Without data, these models and applications are not more than an empty shell containing expert knowledge.

There are many initiatives on Web service chaining, but most are in the field of business applications and only a few in the environmental domain. Moreover, as far as we know, no initiative supports (semi-)automated Web service chaining. Therefore, a research initiative has been set up by a consortium of organisations, excelling in Web service technology and/or in model-based decision support for the environmental domain. This initiative aims at establishing a research project that will focus on combining data services and application services (models and other tools) by web enabling them, making use of existing standards and when necessary, extending the standards.

This initiative, called CHESS, i.e. **ch**aining **e**nvironmental Web **s**ervice**s**, aims at developing a CHESS-system, consisting of the CHESS Integrated Framework (CHESS-IF) and (environmental) Web services (data, models, tools, etc.); see Figure 9. The purpose of the CHESS-system is to enable and support users from the environmental domain to find, bind/chain and execute environmental services distributed over the World Wide Web to create a required application on demand. To realise this the CHESS-IF has to carry out 3 tasks. The first task, performed by the *analyser* with the help of the *meta repository* is to *find* Web services that are required to solve the problem at hand. The second task, performed by the *chain designer*, is to *bind* these Web services to an executable chain of environmental Web services and a workflow, prescribing the chain. In the third task end-users have to *execute* the chain of Web services with the help of a simple user-interface.

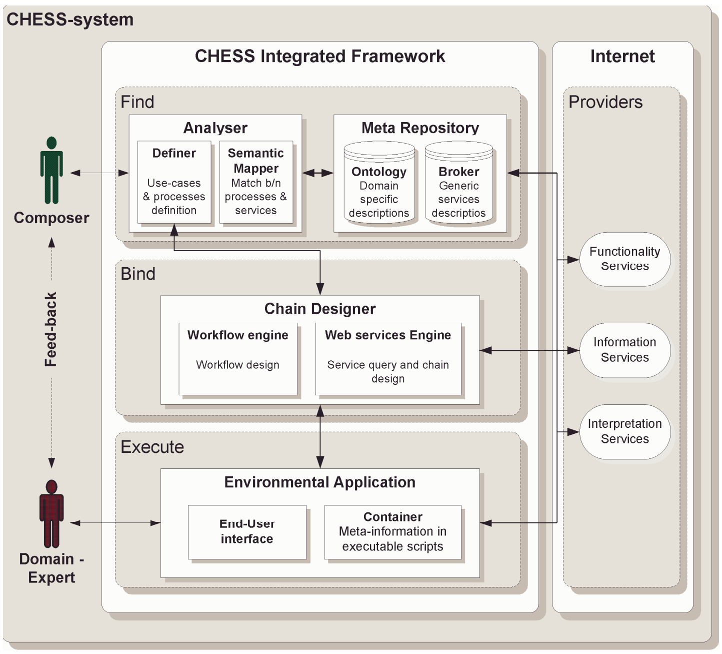

Figure 9: Conceptual overview of the CHESS-system. The focus in this paper is on the definer, while Kassahun et al. (2009) will discuss the semantic mapper, both part of the analyser.

Considerable effort has to be spent to convert a minimum set of Environmental data models, tools, etc. to Web services. The CHESS-system has to be tested in case studies, ranging over a wide spectrum of environmental problem types.

Three different roles of experts can be distinguished. The (human) *composer* interacts with the *definer*-part of the *analyser* to define what functionality, data, etc., is needed to solve the environmental problem at hand. *Domain experts*, i.e. end-users, have to use the resulting *environmental application* to get advice on that problem. *Web service builders* convert existing, relevant data, models, tools, etc. into Web services that are CHESS-compliant.

One of the key factors for the success of the CHESS-initiative will be semantic interoperability between the components of the CHESS-system and between chained environmental Web services. This will be assured by using existing, international and open standards in the *broker*-part of the *meta repository*, e.g. UDDI (Universal Description Discovery and Integration, i.e. a platform-independent registry for businesses worldwide to list themselves on the Internet), WSDL (Web Services Description Language, i.e. an XML-based language that provides a model for describing Web services), OpenMI (Open Modelling Interface and Environment, i.e. a standard for model linkage in the water domain). For the information, missing in these standards, an ontology will be developed. Starting points for this ontology are the work resulting from three EC-sponsored projects OpenMI (in the project Harmon-IT, Gregersen et al., 2005), Seamless (van Ittersum et al., 2008), AquaStress (Kassahun et al., 2008) and the framework discussed by Scholten (2008).

An ontological approach to define the problem and process

Developing the CHESS-system is a challenging exercise. One of the bottlenecks in this ambitious endeavour consists of translating the mental model of some expert in the field (in Chess terms called the composer) into requirements that can be used by the CHESS-IF to find and bind the proper environmental Web services in the right order, generate a workflow depicting this and generate an end-user application (the bound environmental Web services and a web based user interface. The composer's mental model will probably include facts on some specific location and time frame, a suitable mathematical model type, problem domain and many more.

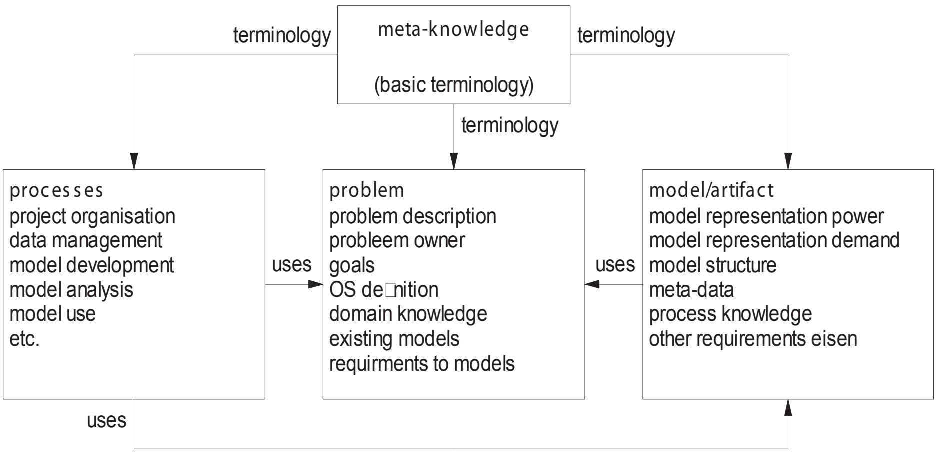

Figure 10: Summary of a structured ontology for multidisciplinary problem-solving, consisting of 4 interrelated ontologies. At a highly abstract level, the ontologies are concepts (presented as rectangles) and (highly abstract) relations between the ontologies (Scholten, 2008).

Scholten (2008) proposed a 4-leaf-clover-like structured ontology for multidisciplinary problem-solving, consisting of a *meta-ontology* (with generic terminology), a *process ontology*, a *problem and object system ontology* and a *model ontology* (or in more generic terms an ontology to describe problem-solving artefacts (see Figure 10). This ontological framework distinguishes layers of increasing specificity (from generic to very detailed). In the context of the project AquaStress (Kassahun et al. 2008) the 4-leaf-clover framework of Scholten (2008) has been extended and implemented for water stress mitigation. This resulted in knowledge bases for the water stress mitigation process (analogous to the process ontology), for the test-sites (analogous to the problem and object system ontology), for options (analogous to the model ontology) and for indicators (linking the options with their effect on the problem and measurable quantities in the object system). In the Seamless project, an ontology has been developed that facilitates communication between models or between model-components, including the units of quantities to be coupled between model modules. To some extent the Seamless results (enabling different models, components and data to let them collaborate as a single (mainly agricultural) simulation model, but unlike what CHESS aims, composite models in Seamless have to run on a single computer and are not a chain of Web services that are only invoked from a single (end-user) computer.

The *definer* part of the *CHESS-analyser* has to handle the rough/vague ideas of the mental model of the *composer* on the problem at hand and provide ways to arrive at a solution or an advice on how to achieve such a solution. The *definer* has to translate this mental model into terms that can be used by the *semantic mapper*. Here we will not present or discuss the design of this *definer component*, but we will merely present some examples of items on which information has to be collected by this interactive input of the CHESS-IF for a specific problem to be solved.

The *composer* has to provide following information and the *definer* will use the terminology from the standards (OpenMI, WSDL, UDDI, etc.) and/or concepts from the additional ontology (part of the meta repository in Figure 1):

- Questions related to problem and object system:
  - ⋆ On problem: problem subject, wanted solution(s), problem complexity, problem description, problem context, problem solving methodology, problem domain(s), etc.
  - ⋆ On object system (OS): OS structure (entities, relations/processes between entities), time frame, location, temporal resolution, spatial resolution, etc.
- On artefact(s): models, data, tools, dynamic, static, proprietary aspects, access issues, etc.
- On problem solving process: what to do (by the chain of environmental Web services), order of things to do, quality of service, etc.

The result is a set of requirements with the views and perceptions of the composer on the problem and its solution at hand in a semi-structured format that can be used by the semantic mapper.

Final remarks

If financed, a CHESS-project can realise a substantial part of an infrastructure to deal with environmental problems, whether induced by climate change or other man-induced environmental impacts. The CHESS-IF can also be used to asses other problems, but in that case the CHESS-system should be extended with adequate Web services required to solve the new kind of problems.

A CHESS-project will set an *ad hoc* standard to environmental Web services. Following this standard will make a Web service CHESS-compliant.

Major risks to deal with include: not enough Web services (first for the test case studies, later to deal with a wide range of environmental problems. A next risk is to end up with an application that does not help the end-user sufficiently. Further risks are related to problems with (semi-)automated chaining, due to technical interoperability or semantic interoperability.

Two topics in CHESS will be innovative: semi-automated chaining of Web services and full semantic interoperability.

Acknowledgements

The ideas in this publication are based on extensive discussions in the context of a FP7-ICT project proposal. The main contributors in these discussions include Klaus Turowski and Sven Overhage (Universität Augsburg, D), Marek Makowski (IIASA, A), Onno Roosenschoon (Alterra, NL), Anker Hojberg (GEUS, DK), Maria-Helena Ramos (Cemagref, F), Lars Chresten Ekebjaerg (DHI, DK), Irina Ribarova and Galina Dimova (UACEG, BG), George Arampatzis (NTUA, GR).

References

[1] Gregersen, J.B., Gijsbers, P.J.A., Westen, S.J.P., Blind, M., 2005. OpenMI: the essential concepts and their implications for legacy software. Advances in Geosciences 4, 37-44.

[2] Kassahun, A., Blind, M., Krause, A., Roosenschoon, O., 2008. Integrated Solution Support System for Water Management. iEMSs 2008: International Congress on Environmental Modelling and Software, Integrating Sciences and Information Technology for Environmental Assessment and Decision Making, 4th Biennial Meeting of iEMSs, `http://www.iemss.org/iemss2008/uploads/W6/PositionPaper.pdf`.

[3] Kassahun, A., Kramer, M.R., Scholten, H., Beulens, A.J.M., 2009. CHESS Integrated Framework: Mapping problem solving tasks to web services. The 22nd International Workshop on Complex Systems Modeling (CSM).

[4] Scholten, Huub. 'Better Modelling Practice: An Ontological Perspective on Multidisciplinary Model-Based Problem-Solving', Wageningen University, 2008.

[5] van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L.R., Andrea E., van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150-165.