International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

Towards a service-oriented e-infrastructure for multidisciplinary environmental research

Ayalew Kassahun[a], Ioannis N. Athanasiadis[b], Andrea E. Rizzoli[b], Arno Krause[c], Huub Scholten[a], Marek Makowski[d], AdrieBeulens[a]
[a]LDI-Logistics Decision & Information Sciences, Wageningen University, The Netherlands,
[b]IDSIA-Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, Lugano, Switzerland,
[c]Alterra, Wageningen University, The Netherlands, [d]IIASA-The International Institute for Applied Systems Analysis, Laxenburg, Austria.

Abstract: Research e-infrastructures are considered to have generic and thematic parts. The generic part provides high-speed networks, grid (large-scale distributed computing) and database systems (digital repositories and data transfer systems) applicable to all research communities irrespective of discipline. Thematic parts are specific deployments of e-infrastructures to support diverse virtual research communities. The needs of a virtual community of multidisciplinary environmental researchers are yet to be investigated. We envisage and argue for an e-infrastructure that will enable environmental researchers to develop environmental models and software entirely out of existing components through loose coupling of diverse digital resources based on the service-oriented architecture. We discuss four specific aspects for consideration for a future e-infrastructure: 1) provision of digital resources (data, models & tools) as web services, 2) dealing with stateless and non-transactional nature of web services using workflow management systems, 3) enabling web service discovery, composition and orchestration through semantic registries, and 4) creating synergy with existing grid infrastructures.

***Keywords**:* e-infrastructure, virtual research communities, multidisciplinary environmental research, service oriented architecture.

1. INTRODUCTION

Europe is making major efforts in developing research e-infrastructures as *"a new environment for academic and industrial research in which 'virtual communities' share, federate and exploit the collective power of European scientific facilities*" [EC 2009]. The European Strategy Forum on Research Infrastructures (ESFRI) roadmap identifies four major components of e-infrastructures: communication networks, distributed grid infrastructures, high performance computing facilities and digital repositories [EC 2008, p. 84]. The first two are considered to be *generic* applicable to all research communities irrespective of discipline. The last two are considered to be *thematic*, i.e. specific deployments of e-infrastructure to support various virtual research communities.

Several scientific communities are making their own thematic roadmaps identifying the needs of their discipline [EC 2008, p. 9]. Within the environmental science domain e-infrastructures for specific thematic areas such as climate change or biodiversity are being implemented. These includes, for instance, satellite and ground based networks of measurement and monitoring stations, but also large scale analytical and modelling tools [EC 2008]. The needs of multidisciplinary environmental researchers with respect to providing easy access to digital resources and providing easy mechanisms for combining these digital resources into useful environmental models and software are yet to be outlined.

Multidisciplinary environmental research brings together researchers from diverse disciplines that include technical, social, economic, political and legal aspects. Such researcher teams use a variety of data and tools [Kassahun et al. 2008]. We argue that multidisciplinary environmental researchers have, due to extreme data and method heterogeneity, a pressing need to an e-infrastructure that facilitates the creation of environmental software systems through service discovery and loose coupling of various digital resources.

The current state-of-the-art information technology (IT) for coupling of diverse digital resources is based on the Service-Oriented Architecture (SOA) concept. In short, the SOA architecture consists of service providers who offer digital resources as web services and advertise their web services using a centralised service broker hosted by a third party. The broker, who holds key information about service providers and their web services, offers service clients a 'yellow pages' like service – an easy interface and mechanism to search and discover web services. Service clients find appropriate web services from the broker and 'couple' them together to create new software applications.

SOA can be considered as an architecture or a blueprint for creating e-infrastructures. To realize an e-infrastructure using the SOA concept, the component parts of the architecture should be made available and 'interoperable'. Specifically, three things are required: (a) digital resources are converted to web services, (b) a centralized service registry and broker is made available, and (c) frameworks and tools that support the process of putting web services together to create new software applications are made available.

This paper summarizes the current state-of-the-art, and outlines how a future environmental modelling and software e-infrastructure for multidisciplinary environmental research can be realized. The remaining part of the paper is organised as follows: In section 2 we discuss the current state-of-the-art in deploying e-infrastructure for multidisciplinary environmental research. In section 3, we describe briefly the SOA concept, in section 4, we propose an e-infrastructure based on loose coupling and the SOA paradigm. We propose a number of additional features to the SOA architecture. We make concluding remarks in section 5.

2. CURRENT STATE-OF-THE-ART

In the past decades, more powerful computing systems were deployed and more complex environmental systems were investigated, leading to more complex software implementations of models. To deal with complexity, researchers and engineers decompose complex problems into smaller problems for which software components can be found or developed. In fact, in environmental science researchers have long been preoccupied with disciplinary approaches that allows them to solve pieces of larger problems from specific disciplinary perspectives. Thus, a variety of appropriate software components are already available. To address complex, multidisciplinary problems, the idea is to "*link*" or "*couple together*" these components into a complex system that deals with the overall problem. Over the years, more and more mature methods of interfacing and coupling diverse types of components (including databases, simulation models and user interaction systems) have been developed [Rizzoli et al. 1998; Argent 2004; Gregersen et al. 2007; van Ittersum et al. 2008]. However, most of these methods were made for stand-alone software implementations, which operate on a single machine and are meant for a single user. The key element for a collaborative research e-infrastructure is virtualization. In such a setting environmental information is treated as a common and scarce asset that needs to be shared among peers [Athanasiadis 2007].

Eventually, environmental information is required to become a virtual resource, and within the environmental research community a number of, albeit uncoordinated, attempts offer parts and pieces of such an e-infrastructure. Specifically, we identify three types of components in a service-oriented e-infrastructure, that are already available: (a) digital resources available to environmental researchers, (b) meta-repositories and service brokers, and (c) frameworks and support tools for coupling of digital resources. We describe each of these briefly below.

2.1 Digital resources

Digital resources, such as data that come from monitoring facilities, have long been the primary focus in deploying research e-infrastructures. The climate data archive of the Intergovernmental Panel on Climate Change (http://www.ipcc-data.org/), Global Environment Outlook data portal of the United Nations (UN) Environment Programme, FAOSTAT (faostat.fao.org), UN data (http://data.un.org) and many national and institutional data archives are good examples. Until now, the main aim has been to make data easily accessible, irrespective of their geographical location. Recent techniques and new e-infrastructures in environmental monitoring have made it possible to acquire tremendous amounts of data.

Increasingly, environmental researchers want to build more comprehensive and complex models, often composed of existing simulation models, whereby the simulation output of one model is an input for another model. Thus, when data is presented to end users or used as input to a simulation model, that doesn't necessarily mean that raw (input) data is used; it can also be processed data or an output from another simulation model. In complex models, composed of many simulation models, the dependency among models can be very tight and cyclic. For instance, *simulation model A* may require intermediate values of *simulation model B*. That requires that the two models have to be executed in tandem.

In addition, nearly all data will undergo some form of processing before it is useful or meaningful to end users or other researchers. For instance, some users may need data at regional level or at national level, while others need only summary data and still others require data to be presented to them as graphs or charts. Thus, what is an input data for one is for the other an output of a complex simulation or data processing. The distinction between the various data sources, be it raw data, outputs from simulation models or data processing software tools can at times be very vague. We thus refer to them collectively as digital resources.

Researchers within the environmental research community identify often three types of digital resources: *data*, *simulation models* and *data processing and visualisation tools*. Data is the necessary 'raw material'. Simulation models and data processing/visualisation tools are what give meaning to data.

The current state-of-the-art with respect to access to data can best be described as ad-hoc. The most common methods of sharing data is apparently either using websites or sending them via e-mail on request. Data is often offered in various formats – spread sheets, database files, text files, etc. There is in general little information as to its quality and as to where and how data is archived. And when data is being publicly accessible it is often not easy to extract relevant information as shown in a recent study [Casagrandi and Guariso 2009].

The situation with respect to access to simulation models is generally better than that of data. Simulation models are often documented using scientific articles and, recently, journals such as Environmental Modelling and Simulation have "software availability" sections [Casagrandi and Guariso 2009]. Like data, simulation models are mostly made available on the web as downloadable software packages. However, configuring and using them remains a challenge.

The third type of digital resources are tools used for processing or visualisation of data. Geographical information systems provide with good examples. Another recent and more promising generic visualisation tool is "ManyEyes" [Viegas et al. 2007]. ManyEyes is a data visualization platform offering a wide array of data representations possibilities.

2.2 Meta repositories

Access to digital resources is facilitated by storing them in digital repositories. Meta repositories, on the other hand, aim at solving the problem of how to find digital resources and are seen as an important component of any e-infrastructure. However, the terms repository and meta-repository are used with diverse and interchangeable meanings. Many efforts that aim to providing scientific data by putting them in repositories also provide meta data facility that enables search and discovery of data. Here a repository means a storage of digital resources and meta-repository means a repository that contains only meta-data about digital resources that are stored or located elsewhere.

Meta repositories are yellow pages like catalogues that help seekers to find digital resources and providers to advertise themselves and their digital resources. In an ideal situation, like in yellow pages, providers of digital resources publish meta-information about themselves and their digital resources on a centralized and well recognized meta-repository. Within the environmental research community the state-of-the-art with respect to the provision of meta-repositories can be described as ad-hoc. That is for instance the case with UN stat, UN data and many national and institutional (meta) repositories. Within the broader IT community, state-of-the-art in providing meta-registries is based on the Universal Description Discovery and Integration (UDDI) specification. UDDI provides a uniform way to describe services and facilitates their discovery.

2.3 Coupling of digital resources

An ever growing demand of authorities for a more realistic models upon which they aim to improve their decisions has led researchers to find ways of combining data, simulation models and tools into a complex but more realistic models. Increased computational possibilities have also made it possible that this ambitions goal can finally be addressed.

Within the environmental research community a significant breakthrough was the development of software frameworks for the integration and linking of environmental models such as the Open Modelling Interface (OpenMI) specification [Gregersen et al. 2007], the Object Modelling System [Kralisch et al. 2005], the Common Component Architecture [Bui et al. 2008], and others. For instance, OpenMI makes it easy to couple models in the hydrological domain and is gaining more and more acceptance within the environmental research community.

3. THE SOA PARADIGM

A widely accepted technique for building large, flexible and loosely coupled systems is Service Oriented Architecture (SOA). SOA is typically used as an integration mechanism. It enables software developers to integrate pieces of software or software functionalities across heterogeneous technological platforms; it enables also integration of software across organizational boundaries. SOA is also viewed as a means of leveraging existing investments in data and software development [Arsanjani et al. 2003; Srinivasan and Treadwell 2005]. In addition, SOA enables a modular approach to software development.

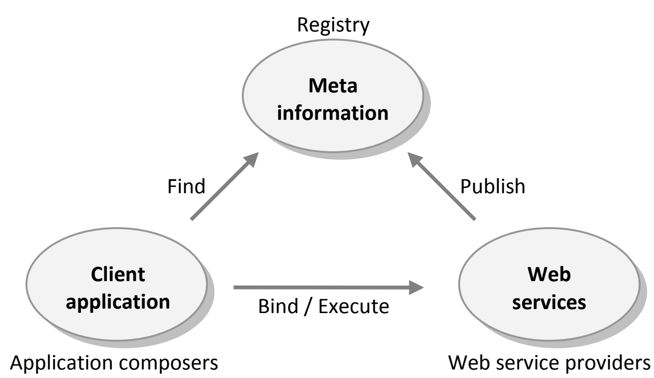

Figure 1. The SOA paradigm

The SOA paradigm provides a blueprint on how to compose easily a new application from existing ones – called web services. SOA is based on distinguishing three distinct roles of software developers and software tools: client applications (application composers), registries or meta repositories (brokers), and web services (web service providers). Service providers offer services; a service broker holds information about services, and service clients browse & find web services using service brokers. Figure 1 depicts how that works.

The SOA approach involves a set of specifications, the most important specifications are *SOAP*, *WSDL* and *UDDI*. The *SOAP (Simple Object Access Protocol)* specification provides a standardized way to send and receive messages over the web between the client application and web services. The *WSDL* (*Web Service Description Language)* specification provides a standardized way of describing what a web service is capable of, including required inputs and produced outputs. The *UDDI* specification provides a standardized way of prescribing how to use a service registry that holds information about service providers and their web services [Curbera et al. 2002].

To realize an e-infrastructure based on the SOA paradigm, a number of factors should be considered. Firstly, the SOA principle requires that digital resources should be made available as web services. This means existing resources (data, models, and tools) should be 'converted' to web services. Since web services have a uniform standard interface they will enable all digital resources to be accessed in the same standard way. Secondly, the web services research is moving towards stateless services, where Representational State Transfer (REST) principles apply [Fielding 2000]. However, almost all non-trivial software systems need to preserve state and transaction information. The statelessness of web services should, thus, be addressed, by using a middleware that preserves state information. Lastly, e-infrastructures are mostly associated with grid computing, and quite a huge investment has already been made to build grid technologies and infrastructures [EGI_DS 2009]. A service oriented e-infrastructure could complement existing grid infrastructures by enabling applications that run on grid platforms to be accessible as web services.

SOA is an area of study within software engineering and it mainly concerns software developers. Researcher teams often build software tools out of necessity for isolated and specific purposes. However, the capacity for modelling complex systems is often beyond the resource capacity of independent research teams. Therefore, the SOA paradigm is being proposed to enable the reuse of existing pieces of software systems with the environmental research community. If complex systems have to be built they can better be built by reusing the works of many others.

4. AN E-INFRASTRUCTURE FOR ENVIROMENTAL RESEARCH

We envisage that future e-infrastructures will be based on software frameworks that will enable researchers to easily and quickly create environmental software applications entirely out of existing components. Researchers without software programming expertise will be able to "drag and drop" icons on screen that represent digital resources and "link" them together to create the desired environmental information or decision support system. There will be no need to 'download' data, models nor software libraries. To 'compose' new environmental applications from existing digital resources no software code will have to be written. The software applications created will be pieces of configuration files and workflow scripts that run in a software framework.

For such frameworks to be realized, an e-infrastructure is required that consists of a wide selection of web services, repositories with rich semantic annotations of web services, workflow middleware's that preserve status and transaction information, and mechanisms for linking to existing grid e-infrastructures.

4.1 Providing environmental digital resources as web services

The first step in building an e-infrastructure based on SOA is to provide digital resources as web services. New digital resources can be designed to become available as web services. Converting existing software or databases to web services is in most cases not a straightforward matter; thus, a systematic approach is required. In literature we find two approaches. In the first approach existing resources are converted to web services manually. Zhang and Yang (2004) identify three methods to convert existing resources to web services, based on the extent to which existing software (sometimes referred to as legacy software) is going to be opened up and modified. They are *white-box*, *black-box* and *mixed* methods. A *white-box* method refers to redesigning or modifying of existing software source code. This method requires often a substantial effort and a good knowledge of the existing system. The main advantage of a white-box method is that the resulting new system will be well suited to the new purpose. A *black-box* process leaves the original software intact and uses “adapters” which are web services whose sole role is to pass messages to and from the existing system. This is sometimes the only option, when original source code is not available, or the existing system is too complex and costly to modify. The third way is a *mixed method*, in which, a partial redesign and implementation of the existing system is required.

As the manual approach is quite expensive, a second approach that aims at simplifying and automating the process is proposed. Towards simplification drives the adoption of widely used interfaces, such as relational database SQL (Structured Query Language) commands, as demonstrated by Amazon’s Relational Database Service (RDS) (http://aws.amazon.com/rds/) and GIS (Geographic Information System) visualization services. Towards automation drives the adoption of rich semantic annotations of the interfaces of existing applications. Semantic annotations of the interfaces enables the generation of software source code for wrapping existing models as demonstrated in Athanasiadis and Janssen [2008]. Semantic techniques are also effective methods of creating new web services by composition from existing web services [Sirin et al. 2002; Rao and Su 2005]. Figure 2 summarizes these approaches.

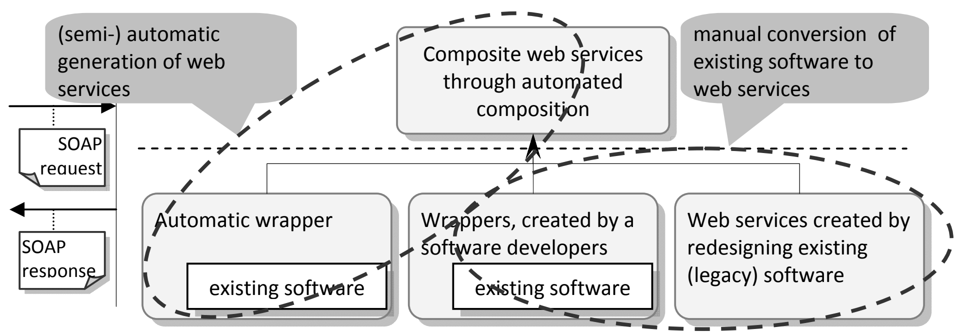

Figure 2. Manual and (semi-) automatic methods of converting existing digital resources into web services.

4.2 Semantic annotations

The second important component of building the future e-infrastructure based on the SOA-paradigm consists of documenting and cataloging of web services in a semantically rich manner. In SOA, web services and their respective publishers are documented according to two specifications, WSDL and UDDI. A WSDL document specifies the interfaces of web services, while a UDDI document specifies information required for cataloguing of web services [Overhage and Thomas 2009]. While publishing meta-information to SOA registries can easily be automated, and indeed is often automated, finding the right web service is essentially a manual activity. The client application developer (application composer) should do the search and make the judgement if the web services found are the right ones for the purpose. Once the developer decided that a web service has been found then its WSDL document is used to generate source code that can be used in the client application.

The idea of semantic annotation is a key part of the creation of the visionary semantic web. The semantic web vision is to define and link information on the web in such a way that computers can deduce meaning and use the information autonomously. A richer documentation using semantic annotations of web services means advanced service discovery, orchestration and composition will be a possibility.

The preferred method of creating semantic annotations is using ontologies. Ontologies have been used to model declaratively the interfaces of components and sometimes the actual logics from which software code is generated in anatomised process [Athanasiadis and Janssen 2008; Rizzoli et al. 2008]. Standard specifications such as UDDI can also be seen as a means of semantic annotation. For instance, Overhage and Thomas [2009] proposed a number of improvements to UDDI specification to enhance the description of web services and their providers thereby to improve web service discovery and configuration.

The problem of using specifications as a mechanism for semantic annotation is that both the creation and improvement processes of specifications are very slow and tedious. Ontologies on the other hand can be developed relatively quickly and have an added advantage of being verifiable or provable. An important advantage of specifications is that specifications, especially those related to the semantic web, are well recognized and at the same time organized at different abstraction levels for reuse. Many ontologies, on the other hand, lack wide spread recognition and are less frequently reused. An approach often used to promote reuse is to organise ontologies based on layers, from generic to specific. For instance, Scholten [2008] proposes an ontological layer structure for three areas relevant for multidisciplinary environmental research, which include: ontologies for problems and their object systems, ontologies for processes including modelling, and ontologies for models and other problem-solving artefacts.

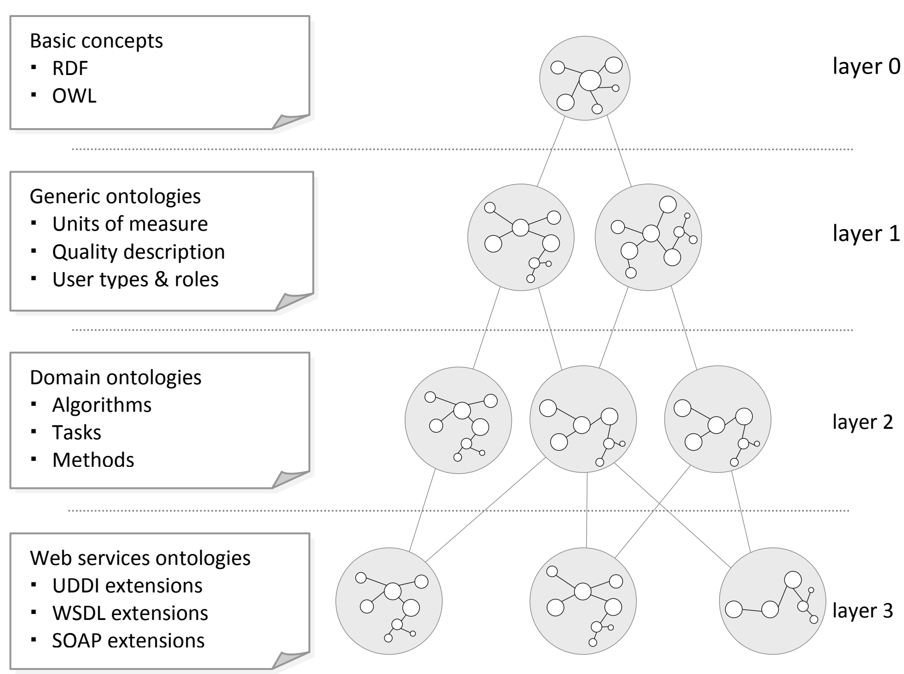

Figure 3. Ontologies that can be used to semantically annotate web services and their interfaces organized in layers of reusability.

Various ontologies can be used to semantically annotate SOA specifications. The aim of semantic annotation is to enable environmental researchers describe the desired environmental software as a series of interrelated tasks which then can be mapped to suitable web services. To achieve this, various SOA documents (i.e. SOAP, WSDL and UDDI documents) associated with web services need to be semantically annotated. To organise and reuse ontologies that can be used inside SOA specifications we propose a class of ontologies organised in layers as shown in Figure 3. Layer 0 consists of basic concepts used in any ontology. Layer 1 contains generic ontologies relevant to any software framework, such as measurement units ontology or quality ontology codifying the

ISO ISO/IEC 9126 quality model. Layer 2 consists of domain ontologies of a given domain of application, in our case the environmental domain, such as ontologies for modelling processes. The last layer – layer 3 – consists of specific extensions relevant for each specification used in the SOA architecture such as those recommended extensions by Overhage and Thomas [2009].

4.3 Workflow management

Web services are stateless and ways have to be found to preserve state information between invocations. The current approach to solve this problem is using workflow management systems. Currently there are two approaches to workflow management: Scientific Workflows and Business Process Modelling (BPM). Though both scientific workflows and business process modelling approaches aim at modelling and execution of workflows, both have evolved along different paths. Comparative studies show that the main reasons for these different approaches are the data intensive nature and the need for experimentation in scientific workflows, while the focus in business processes lies in supporting various users with differing roles [Barga and Gannon 2007; Ludäscher et al. 2009; Yildiz et al. 2009]. The need for supporting various users with differing roles is indeed vital in multidisciplinary research because different researchers with diverse expertise and roles are involved. An example of this last is the Modelling Support Tool (MoST) developed to support multidisciplinary modelling [Scholten et al. 2007]. However, MoST supports only manual activities and there are no facilities for web service invocation or automation of tasks. Recently there are promising efforts in using broadly accepted standards for business process modelling such as BPEL (Business Process Execution Language) for scientific workflows [Candela et al. 2007; Clementi et al. 2008]. An advantage of adopting BPM standards is the widespread support from the software industry to the standards, and as a consequence the availability of industrial strength software implementations [Barker and van Hemert 2008]. Using BPM for scientific workflows approach is not without shortcomings as outlined by Ludäscher et al. [2009]. These shortcomings are more critical for research in which computational experiments are conducted and few human interventions are required. We believe that in large multidisciplinary research the stated shortcoming are less critical or even not shortcomings at all. We agree, however, with Ludäscher et al. that further research for "cross-fertilization" of the two approaches is beneficial.

4.4 Link to grid e-infrastructures

An SOA based e-infrastructure is in principle complementary to a grid based e-infrastructure. While an SOA-approach solves the problems of the service client (whether it is for accessing data or computing facility) by making remote access to software possible, a grid-approach solves performance problems service providers may have by enabling them to run their applications on multiple computers. Combining both approaches (i.e. by making grid-based applications accessible as web services) will result in a synergy that is important, especially when high performance computation is required by some web services. Good examples in which SOA is used as 'top layer' for the purpose of combining distributed services while the grid approach is used as a 'lower layer' for the purpose of sharing computational resources are the CYCLOPS [Mazzetti et al. 2009] and DILIGENT [Candela et al. 2007] software systems.

5. CONCLUSION

In this paper, we have reviewed fragmentary developments in e-infrastructure technology for multidisciplinary environmental research. Though our review is not comprehensive, it shows the need for a systematic approach. We envisage an e-infrastructure that will enable multidisciplinary teams to compose environmental software applications, whether they are models, information systems or decision support systems, from existing digital resources (data and software) using the SOA architecture. We identify four areas where work is still required. The first is providing digital resources as web services using different approaches, from re-designing existing software to automated conversion of existing resources to web services. The second is semantic annotations of web services to enable service discovery as well as the creation of new web services through automated composition of web services. We propose augmenting existing SOA specifications with ontologies and a layered approach to organising various ontologies to promote the reuse of existing ontologies. The third area of focus is on managing state information since a number of interdependent web services are used in an application. We propose the adaption of widely accepted standards for business process management that focuses on supporting diverse users in a workflow. Finally, any future e-infrastructure based on the SOA paradigm need to exploit the already existing grid infrastructures. This last is particularly necessary since many large scale environmental models may need to run on grid e-infrastructures for performance reasons. These grid-based software applications need to be made available as web services in order to create the synergy of high performance of grid infrastructures with loose coupling and easy access of web services.

ACKNOWLEDGMENTS

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 Onno Roosenschoon and Sander Jansen (Alterra, The Netherlands), Klaus Turowski and Sven Overhage (University of Augsburg, Germany), Anker Højberg (GEUS, Denmark), Franz Thomsen (DHI, Denmark), Irina Ribarova and Galina Dimova (UACEG, Bulgaria), George Arampatzis (NTUA, Greece), Michael Flechsig and Markus Wrobel (PIK, Germany) and Bert Jagers (Deltares, The Netherlands).

REFERENCES

Argent, R. M., An overview of model integration for environmental applications--components, frameworks and semantics, *Environmental Modelling & Software* 19(3), 219-234, 2004.

Arsanjani, A., B. Hailpern, J. Martin and P. Tarr, Web services: Promises and compromises, *ACM Queue* 1(1), 48-58, 2003.

Athanasiadis, I. N., Towards a Virtual Enterprise Architecture for the Environmental Sector, *Agent and Web Service Technologies in Virtual Enterprises*, 256, 2007.

Athanasiadis, I. N. and S. Janssen, Semantic mediation for environmental model components integration, *Information Technologies in Environmental Engineering* 1, 3-11, 2008.

Barga, R. and D. Gannon, Scientific versus Business Workflows. *Workflows for e-Science*, 9-16, 2007.

Barker, A. and J. van Hemert, Scientific Workflow: A Survey and Research Directions. *Parallel Processing and Applied Mathematics*, 746-753, 2008.

Bui, V., B. Norris, K. Huck, L. C. McInnes, L. Li, O. Hernandez and B. Chapman, A component infrastructure for performance and power modeling of parallel scientific applications. Proceeding of the Component Based High Performance Workshop - CompFrame/HPC-GECO Workshop Series, CBHPC'08, art. no. 6, ACM, 2008.

Candela, L., F. Akal, H. Avancini, D. Castelli, L. Fusco, V. Guidetti, C. Langguth, A. Manzi, P. Pagano, H. Schuldt, M. Simi, M. Springmann and L. Voicu, DILIGENT: integrating digital library and Grid technologies for a new Earth observation research infrastructure, *International Journal on Digital Libraries* 7(1), 59-80, 2007.

Casagrandi, R. and G. Guariso, Impact of ICT in Environmental Sciences: A citation analysis 1990–2007, *Environmental Modelling and Software* 24(7), 865-871, 2009.

Clementi, L., C. Cacciari, M. Melato, R. Menday and B. Hagemeier, A Business-Oriented Grid Workflow Management System. *Euro-Par 2007 Workshops: Parallel Processing*, 131-140, 2008.

Curbera, F., M. Duftler, R. Khalaf, W. Nagy, N. Mukhi and S. Weerawarana, Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI, *IEEE Internet computing*, 86-93, 2002.

EC, European Strategy Forum on Research Infrastructures (ESFRI). Roadmap report 2008. ftp://ftp.cordis.europa.eu/pub/esfri/docs/esfri_roadmap_2008_update_20090123.pdf, 2008.

EC, CORDIS: ICT programme: e-Infrastructure overview, http://cordis.europa.eu/fp7/ict/e-infrastructure/overview_en.html , accessed on 27 November 2009., 2009.

EGI_DS, European Grid Initiative »Towards a sustainable production grid infrastructure« Retrieved 27 November.

Fielding, R. T., Architectural styles and the design of network-based software architectures, Citeseer, 2000.

Gregersen, J. B., P. J. A. Gijsbers and S. J. P. Westen, OpenMI: Open modelling interface, *Journal of hydroinformatics(Print)* 9(3), 175-191, 2007.

Kassahun, A., M. Blind, A. Krause and O. Roosenschoon, Integrated solution support systems for water management. iEMSs 2008, International Congress on Environmental Modelling and Software Barcelona, 2008.

Kralisch, S., P. Krause and O. David, Using the object modeling system for hydrological model development and application, *Advances in Geosciences* 4, 75-81, 2005.

Ludäscher, B., M. Weske, T. McPhillips and S. Bowers, Scientific Workflows: Business as Usual? *Business Process Management*, 31-47, 2009.

Mazzetti, P., S. Nativi, V. Angelini, M. Verlato and P. Fiorucci, A Grid platform for the European Civil Protection e-Infrastructure: the Forest Fires use scenario, *Earth Science Informatics* 2(1), 53-62, 2009.

Overhage, S. and P. Thomas, WS-Specification: Specifying Web Services Using UDDI Improvements. *Web, Web-Services, and Database Systems*, 100-119, 2009.

Rao, J. and X. Su, A Survey of Automated Web Service Composition Methods. *Semantic Web Services and Web Process Composition*, 43-54, 2005.

Rizzoli, A. E., J. R. Davis and D. J. Abel, Model and data integration and re-use in environmental decision support systems, *Decision Support Systems* 24(2), 127-144, 1998.

Rizzoli, A. E., M. Donatelli, I. N. Athanasiadis, F. Villa and D. Huber, Semantic links in integrated modelling frameworks, *Mathematics and Computers in Simulation* 78(2-3), 412-423, 2008.

Scholten, H., Better modelling practice: an ontological perspective on multidisciplinary model-based problem-solving. Wageningen, Wageningen University. PhD Thesis, 314, 2008.

Scholten, H., A. Kassahun, J. C. Refsgaard, T. Kargas, C. Gavardinas and A. J. M. Beulens, A methodology to support multidisciplinary model-based water management, *Environmental Modelling & Software* 22(5), 743-759, 2007.

Sirin, E., J. Hendler and B. Parsia, Semi-automatic Composition of Web Services using Semantic Descriptions, Citeseer, 2002.

Srinivasan, L. and J. Treadwell, An overview of service-oriented architecture, web services and grid computing, *HP Software Global Business Unit*, 2005.

van Ittersum, M. K., F. Ewert, T. Heckelei, J. Wery, J. Alkan Olsson, E. Andersen, I. Bezlepkina, F. Brouwer, M. Donatelli, G. Flichman, L. Olsson, A. E. Rizzoli, T. van der Wal, J. E. Wien and J. Wolf, Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS), *Agricultural Systems* 96(1-3), 150-165, 2008.

Viegas, F. B., M. Wattenberg, F. Van Ham, J. Kriss and M. McKeon, Manyeyes: a site for visualization at internet scale, *IEEE Transactions on Visualization and Computer Graphics* 13(6), 1121, 2007.

Yildiz, U., A. Guabtni and A. H. H. Ngu, Business versus Scientific Workflows: A Comparative Study, IEEE Computer Society, 2009.