

Orchestration of Geospatial Processes with RichWPS – a Practical Demonstration

Dennis Ziegenhagen¹, Jörn Kohlus², Rainer Roosmann¹

Abstract

Distributed geospatial data and services can be used to form Spatial Data Infrastructures (SDI). Current technological developments enable SDI to process large amounts of geospatial data. In this context, Web Processing Services (WPS) can be used as an open interface standard for accessing and executing geospatial processes. Since many SDI are based on Service-Oriented Architectures (SOA), the orchestration of services is enabled by composing existing services. The results are higher services, e.g. complex geospatial applications. The RichWPS research project focuses on analysing and providing practical approaches of web service orchestration. For this, software components are developed to build a modular orchestration environment. Real-life examples of geospatial processing are analysed and divided into simpler, reusable sub-processes. Using the RichWPS environment, the sub-processes are implemented and arranged as workflow models, which are deployed to an extended WPS server and interpreted by its orchestration engine. This enables workflows to be executed by WPS clients. This paper presents the RichWPS environment and one of the analysed real-life scenarios. The example scenario “Macrophyte Assessment” contains the assessment of water plants in order to report on the water’s qualitative status. This meets requirements of the Water Framework Directive (WFD) and the Marine Strategy Framework Directive (MSFD) of the European Union. The scenario is analysed in order to identify elementary processes. With the use of RichWPS components these processes are orchestrated to form an executable assessment procedure.

1. Introduction

With the current developments in Geographic Information Systems (GIS), Spatial Data Infrastructures (SDI) offer opportunities to distribute and process spatial data [1]. SDI implementations are mainly based on the Service-Oriented Architecture paradigm (SOA) and offer access to shared geospatial data and applications [2]. Therefore interoperability and a uniform service access are necessary. Both aspects are supported by open web service standards. In the scope of geospatial services, the Open Geospatial Consortium (OGC) defines widely used standards, commonly referred to as OGC Web Services (OWS).

This paper focuses on OGC’s Web Processing Services (WPS) to enable a modular orchestration of distributed geospatial processes and data. In terms of the SOA paradigm, *service orchestration* is a method to design higher services by composing existing services [3]. As the web services technology has matured in the past years [4], it meets the requirements of distributed geospatial systems. Thus, web services are able to form the technological basis of geospatial infrastructures and enable further research on web service application and opportunities for future extensions.

1.1. The RichWPS Research Project

RichWPS is a research project funded by Germany’s Federal Ministry of Education and Research³. It focuses on inspecting and improving the practical use of WPS inside distributed geospatial

¹ Osnabrück University of Applied Sciences - Faculty of Engineering and Computer Science, Albrechtstr. 30, 49076 Osnabrück, Germany, {d.ziegenhagen, r.roosmann}@hs-osnabrueck.de

² Schleswig-Holstein’s Government-Owned Company for Coastal Protection, National Parks and Ocean Protection, Schlossgarten 1, 25832 Tönning, Germany, joern.kohlus@lkn.landsh.de

systems like Spatial Data Infrastructures (SDI). Research results are compared to existing solutions, e.g. web service orchestration using the Business Process Execution Language (BPEL) together with the Web Services Description Language (WSDL).

To achieve and analyse the desired project results, multiple software components are developed which together form a comprehensive OWS orchestration environment. The development is done by Osnabrück University of Applied Sciences⁴ and Disy Informationssysteme GmbH⁵. The environment includes a WPS 1.0.0 implementation conforming to OGC's standard. To increase the practical functionality, the implementation also picks up ideas of transactional WPS (WPS-T) and other aspects of the forthcoming standard version WPS 2.0. The aim is to enable an easy creation and use of complex geospatial processing applications for a large domain of users.

Real life use cases and example scenarios are provided by the project partners, Schleswig-Holstein's Government-Owned Company for Coastal Protection, National Parks and Ocean Protection⁶ and Germany's Federal Waterways Engineering and Research Institute⁷. The scenarios contain applications of geospatial processing with the potential to be described as a workflow arrangement of less complex processes. Hence, the initial scenario processes are divided into functional units and each of these units is made available as a WPS. The aim is to extract simple processes which can be utilised as widely reusable building blocks for geospatial applications.

With the RichWPS environment, data and services like the described processing units can be arranged to form workflow models. Loops, conditional forks and other control structures allow dynamic and flexible workflows. Arranged models can be validated, tested and finally deployed onto a server which enables workflows to be executed by WPS client applications.

2. The RichWPS environment

To build a flexible and modular orchestration environment, three main software components and a domain specific language (DSL) are developed. Creating, testing and deploying workflow models are tasks of the *RichWPS ModelBuilder*. It uses the *RichWPS SemanticProxy* as a directory service to look up and receive semantically enriched information about available OWS. Validated workflow models are translated into the *RichWPS Orchestration Language* (ROLA). This DSL is used to transfer workflow descriptions to the *RichWPS Server*, which interprets the ROLA document and uses it to set up the orchestration engine. After a workflow has been deployed to the server as a new process, it is able to be executed via the server's WPS interface. A visual representation of the environment, including the connected components, is shown in Figure 1.

2.1. ModelBuilder

With the ModelBuilder, the task of service orchestration is abstracted and simplified in a graphical and interactive way. Creating a model of the desired workflow hides technical orchestration details and enables editing and adjusting workflows without directly applying changes on a deeper level (e.g. to the server or the orchestration engine). Workflow models are also easy to share with other users and can be used as a base for new workflows to build on.

³ Bundesministerium für Bildung und Forschung, <http://www.bmbf.de>

⁴ <http://www.hs-osnabrueck.de/>

⁵ <http://www.disy.net/>

⁶ Landesbetrieb für Küstenschutz, Nationalpark und Meeresschutz des Landes Schleswig-Holstein, <http://www.lkn.schleswig-holstein.de>

⁷ Bundesanstalt für Wasserbau, <http://www.baw.de>

Target users of the core modelling tool are experts in geospatial domains. Therefore the user interface and provided functionality should be intuitive and easy to understand. Knowledge of technical details and the underlying systems is not necessarily needed to create complex processing workflows. A simple and well-structured interface design supports the usability. Dynamic visual hints and textual messages provide helpful feedback without distracting users from their tasks.

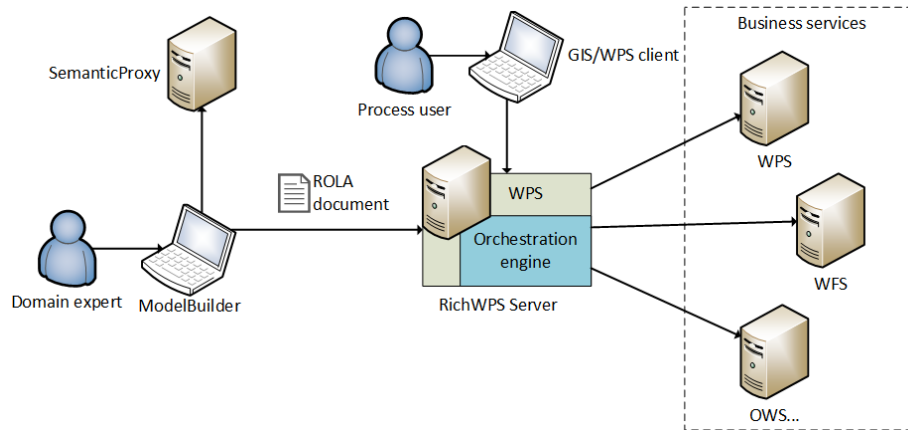


Figure 1 RichWPS environment

The ModelBuilder's main window (Figure 2) is divided into four areas. The tree view on the left contains a hierarchically structured list of available modelling elements like services and data sources. Modelling elements can be added to the editor in the window centre via drag and drop gestures. Using the mouse, compatible input ports (green) and output ports (blue) of modelling elements can be connected. Single characters inside input and output ports represent the used data types which are defined by the WPS standard [5]. For example, "L" is *LiteralData* and "C" represents *ComplexData*.

All elements, including connection lines, can be moved and arranged as needed. Additionally, an auto-layout mechanism is provided. The example model shown in Figure 2 also contains explicit input ports (yellow boxes at the top) and output ports (black box at the bottom) which will be the inputs and output of the resulting WPS after deploying the workflow onto the server.

Further information of selected modelling elements is shown in the properties view on the right side of the window. Users are also able to enter necessary input data in the properties view, e.g. literal numbers or strings.

Status information, error messages and other hints are displayed and organized in multiple tabs at the window's bottom. Each tab contains messages belonging to a specific component: SemanticProxy, model editor or RichWPS server. New and unread messages are highlighted.

Other functionality like workflow deployment and process lifecycle management is accessible via the window's menus and toolbar. Helping information is consistently provided for all GUI elements and enables users to easily find the desired function.

In addition to the modelling elements shown in the example, loop and control structures can be used to create more powerful and dynamic workflows. For example, conditional forking elements allow the execution of processes only if a given condition is true (e.g. if specific geometries intersect). This is comparable to *if*-statements in programming languages. Accordingly, an *else*-fork of the condition exists, in order to execute other processes if the condition is not true.

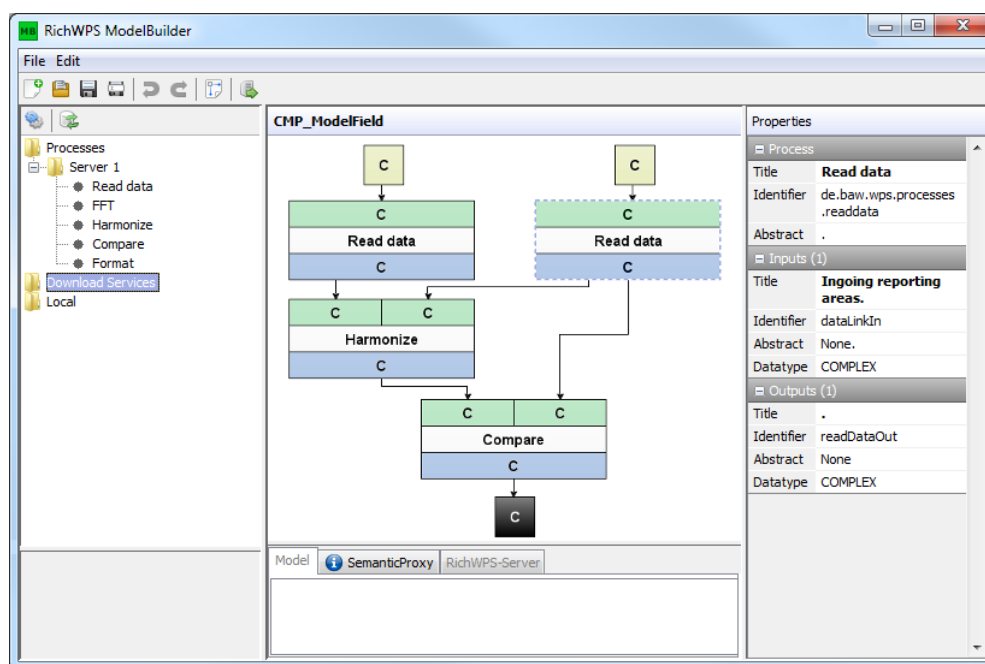


Figure 2 RichWPS ModelBuilder

2.2. Semantic Proxy

The ModelBuilder receives information about available OWS by connecting to the SemanticProxy. Hence, the SemanticProxy is designed as a directory service which can be searched via a RESTful interface. Information about available OWS is stored in a database and follows the Resource Description Framework (RDF) specification [6].

RDF enables a simple but flexible enrichment of service descriptions with semantic information. For database interaction and modification small client applications are implemented. Their functionality is also integrated into the ModelBuilder to create a bundled core application inside the RichWPS environment.

2.3. RichWPS Server

Being based on the 52°North WPS server⁸, the RichWPS Server's WPS implementation is extended by a custom orchestration engine. The server's WPS-T interface enables the ModelBuilder and other client applications to deploy workflow descriptions via ROLA documents. These are interpreted by the orchestration engine in order to offer a centralised, controlled execution of the workflow and the involved processes and services.

Additionally, the server provides extended mechanisms for client communication and interaction. They allow ModelBuilder instances to monitor workflow executions and to receive details on the execution status and occurring errors. It is also possible to pause or stop running workflow executions.

The combination of the ModelBuilder and the RichWPS Server forms a comprehensive setup to test and debug workflow models. Test results can be used to improve workflows and to decide on which processes are to be used or preferred.

⁸ <http://52north.org/communities/geoprocessing/wps/index.html>

2.4. RichWPS Orchestration Language ROLA

ROLA documents are transferred from the ModelBuilder to the RichWPS Server and contain a textual translation of the workflow. They can also be considered as a pre-processing towards the technical orchestration. Thus, ROLA is a domain specific language having a strict syntax that can be interpreted by server components. But it is designed to be still human readable and understandable by choosing appropriate keywords and grammar rules.

3. Example Scenario: Macrophyte Assessment

Macrophytes are aquatic plants growing in or near water. The assessment of macrophyte occurrences is one of the real life use cases which are used to analyse and proof the RichWPS concept. The following section describes the use case itself and its implementation using the RichWPS environment.

3.1. Scenario Description

The Water Framework Directive (WFD) [7] instructs European Union (EU) member states to achieve good quality status of waters. States are directed to assess and report ecological quality values of rivers and seas. Another important directive for the marine sector is the Marine Strategy Framework Directive (MSFD) [8]. Its main objective is to achieve the Good Environmental Status (GES) of waters by observing and controlling improvements of the marine environment's quality.

Frequent quality measurements and reports are required by both, the WFD and the MSFD. The EU demands automated reporting systems which deliver the results via web services to the European Commission. This meets requirements of the INSPIRE directive [9] and includes processing large amounts of data in the fields of chemical and biological quality components.

Implementing WFD and MSFD instructions for Schleswig Holstein's Wadden Sea is inside the scope of Schleswig-Holstein's Government-Owned Company for Coastal Protection, National Parks and Ocean Protection. This includes the reporting areas "Northern Frisia" and "Dithmarschen". The presented example scenario covers the creation of macrophyte assessment reports for these areas.

The occurrence and the percentage of the macrophytes *seagrass* and *green macroalgae* is considered as an indicator for the quality of waters [10]. As a simplified summarisation, the quality status of the Wadden Sea is "good" if it contains more seagrass and less macroalgae. The assessment is based on five quality parameters, each evaluated within one of five categories that range from "0 - bad" to "4 - very good". Combined with the categories, the parameters form an assessment matrix which is used to get the resulting quality status based on the evaluated parameters [11]. The final report contains assessment status for the relevant years: a specific reporting year and additionally the past six years.

3.2. Processes and Orchestration

An implementation of the described assessment process as a single WPS has already been done by Wössner [12]. The evaluation process is a chain of filtering and GIS operations. The technical environment for analysing and reporting is defined through the rules of the Marine Spatial Data Infrastructure (MDI-DE), a SDI [13].

In the presented example scenario, this process is divided into reusable and less complex sub-processes. For the assessment, specific input data is necessary, for example the year of assessment, reporting area geometries and raw data of macrophyte occurrences according to the MSFD.

Table 1 contains the resulting sub-processes after splitting up the functionality of the existing assessment process. Each process is implemented as a single WPS, indicated by the green data input ports at the top and the blue output ports at the bottom.

The workflow model shown in Figure 3 uses the processes *SelectReportingArea* and *Intersect* multiple times to demonstrate the desired reusability. Executing the workflow delivers assessed quality status for the reporting areas as ComplexData.

Process	Description
<div>C</div> <div>L</div> <div>SelectReportingArea</div> <div>C</div>	Returns the reporting area geometry which corresponds to a literal input ID (e.g. "NF" for Northern Frisia)
<div>C</div> <div>L</div> <div>MSRLD5selection</div> <div>C</div> <div>C</div> <div>C</div>	Gets raw data for algae, seagrass and relevant years from existing MSFD data which is relevant for the given year ⁹
<div>C</div> <div>C</div> <div>SelectTopography</div> <div>C</div> <div>C</div> <div>C</div>	Selects topography based on existing topographies and relevant years
<div>C</div> <div>C</div> <div>C</div> <div>Intersect</div> <div>C</div>	Returns the intersection of reporting areas, topographies and relevant years
<div>C</div> <div>C</div> <div>C</div> <div>C</div> <div>C</div> <div>C</div> <div>Characteristics</div> <div>C</div>	Returns the reporting area assessments based on the input data

Table 1 Example Scenario Processes

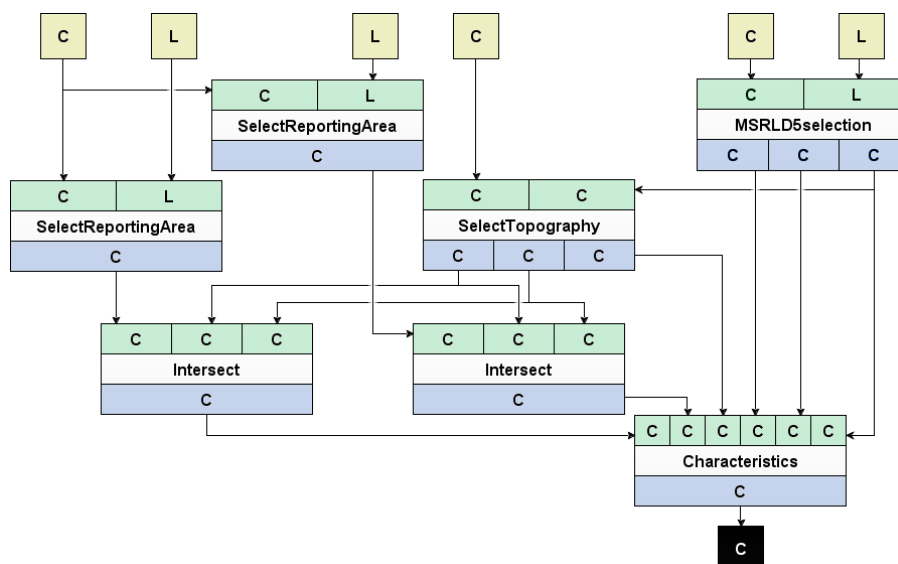


Figure 3 The Assessment Process as a Workflow Model

After implementing the single processes as WPS, information about the processes is stored and provided by the SemanticProxy. Using the ModelBuilder the processes are connected and arranged to model the workflow as shown in Figure 3. The modelling result is translated to ROLA and deployed onto a RichWPS Server. After the successful deployment the server's orchestration

⁹ MSRL (Meeresstrategie-Rahmenrichtlinie) is the German translation of MSFD, D5 means descriptor 5 which is defined by the MSFD

engine allows executing the workflow by one single WPS *execute* call. This also enables a black box view on the workflow as a covering process as shown in Figure 4. Because of the server's implementation of the WPS 1.0.0 standard, execution calls can be done by any WPS client. The new process is also communicated to the SemanticProxy and thus can be used as a modelling element for other workflow models.

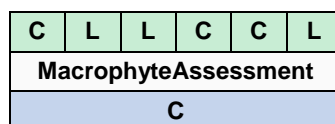


Figure 4 Black box View of the deployed Workflow

4. Related Work

Web service orchestration inside SOA-based infrastructures is discussed in [14]. At first, the paper summarises relevant technologies, requirements and opportunities. It then focuses on increasing the interoperability of published workflows by introducing the WPS-T standard which the RichWPS project uses.

Re-using distributed geospatial web services to form a business logic component has been demonstrated in [15]. It offers an interface between users and the geospatial data inside SOA infrastructures. Additionally, applications of the WPS standard are demonstrated.

Orchestrating web services using BPEL has been demonstrated in [16]. Mentioned are the advantages of using graphical tools over a hard-coded service orchestration. The use of WSDL is necessary to bring web services together with BPEL. The article states this circumstance's potential of danger to BPEL-processes (and thus to the whole orchestration).

Different software tools for graphical workflow modelling are available. They either focus on a specific subject or offer a comprehensive functional range as Taverna Workbench¹⁰ does. An example for WPS orchestration using Taverna Workbench can be found in [17]. In the example, WPS instances require WSDL descriptions and, instead of the often chosen BPEL, the Simple Conceptual Unified Flow Language (SCUFL) is used.

5. Conclusion and Outlook

In early project stages, comprehensive researches and analyses have been done to gather the technical state of relevant technologies and existing approaches. Thus, the project's scope and basic conditions could be specified in detail to precisely prepare the software engineering.

At the time of writing, the software development is advanced and integrating the components to form the environment has been tested. During the development, the example scenarios have partly been implemented to test the components. First results proof the RichWPS environment approach and show the practical usability.

Currently a main task is to continue the integration in order to completely implement the scenarios. This will allow a closer research to get the desired project's results.

6. Acknowledgement

This work has been funded by the German Ministry of Education and Research. The authors also would like to thank the project partners, the Federal Waterways Engineering and Research Institute, and Disy Informationssysteme GmbH for their support.

¹⁰ <http://www.taverna.org.uk/>

References

- [1] T. Wehrmann, S. Gebhardt, V. Klinger and C. Künzer, “Data processing using Web Processing Service orchestration within a Spatial Data Infrastructure“, in *Proceedings of the 34th International Symposium on Remote Sensing of Environment*, 2011.
- [2] T. Foerster, B. Schäffer, B. Baranski and J. Brauner, “Geospatial Web Services for Distributed Processing - Applications and Scenarios“, in *Geospatial Web Services: Advances in Information Interoperability*, 2011, pp. 245-286.
- [3] N. Josuttis, *SOA in Practice*, Sebastopol: O'Reilly & Associates, 2007.
- [4] P. Zhao, T. Foerster and P. Yue, “The Geoprocessing Web“, in *Computers & Geosciences* 47, Elsevier, 2012, pp. 3-12.
- [5] Open Geospatial Consortium, “OpenGIS Web Processing Service Version 1.0.0“, 2007. [Online]. Available: <http://www.opengeospatial.org/standards/wps>.
- [6] W3C, “Resource Description Framework (RDF) Model and Syntax Specification“, 1999. [Online]. Available: <http://www.w3.org/TR/PR-rdf-syntax/>.
- [7] The European Parliament and of the Council of the European Union, Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, 2000.
- [8] The European Parliament and of the Council of the European Union, Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive), 2008.
- [9] The European Parliament and of the Council of the European Union, Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007.
- [10] K. Reise and J. Kohlus, “Seagrass, an indicator goes astray“, in *Observing the Coastal Sea. An Atlas of Advanced Monitoring Techniques.*, Geesthacht, 2008, pp. 64-67.
- [11] T. Dolch, C. Buschbaum and K. Reise, “Seegrass-Monitoring im Schleswig-Holsteinischen Wattenmeer 2008“, 2009.
- [12] R. Woessner, “Untersuchungen zur praktischen Nutzbarkeit des OGC Web Processing Service (WPS) Standards (Masterthesis)“, Hochschule Karlsruhe Technik und Wirtschaft, 2013.
- [13] J. Kohlus and H.-C. Reimers, “Neue Herausforderungen im Datenmanagement für das europäische Meeresmonitoring - Das Projekt MDI-DE - Marine Daten-Infrastruktur in Deutschland“, in *From Brazil to Thailand - New Results in Coastal Research. Coastline Reports (16)*, Rostock, EUCC - Die Küsten Union Deutschland e.V., 2010, pp. 115-126.
- [14] B. Schäffer, “Towards a Transactional Web Processing Service (WPS-T)“, in *GI Days 2008 - Proceedings of the 6th Geographic Information Days*, 2008.
- [15] C. Kiehle, “Business logic for geoprocessing of distributed geodata“, in *Computers & Geosciences* 32, Elsevier, 2006, pp. 1746-1757.
- [16] A. Weiser and A. Zipf, “Web Service Orchestration of OGC Web Services“, in *Geomatics Solutions for Disaster Management*, Berlin, Springer Verlag, 2007, pp. 239-254.
- [17] J. de Jesus, P. Walker, M. Grant and S. Groom, “WPS orchestration using the Taverna workbench: The eScience approach“, in *Computers & Geosciences* 47, 2012, pp. 75-86.