Geoprocessing Services for Spatial Decision Support in the Domain of Housing Market Analyses

Experiences from Applying the OGC Web Processing Service Interface in Practice

Beate Stollberg[1] and Alexander Zipf[1]

[1] Department of Geography,
University of Bonn,
Meckenheimer Allee 172, Bonn, Germany
{stollberg, zipf}@geographie.uni-bonn.de

Abstract. Housing market analysts typically rely on *Geographic Information Systems* (GIS). The new *Web Processing Service* (WPS) standard proposed by the *Open Geospatial Consortium* (OGC) offers an alternative that supports desktop GIS analysis functionalities over the Internet via a standardised interface. This article presents how the WPS specification was applied for housing market analyses and demonstrates the flexibility of the GIS functionalities that can be used in specific scenarios while at the same time providing them as WPS processes. Furthermore, the specification was also used for chaining Web Services and demonstrates that it is a viable alternative to the usage of established standards for *Web Service Orchestration* (WSO) such as the *Business Process Execution Language* (BPEL). The advantages and disadvantages of such approaches are discussed.

Keywords: geoprocessing, Web Processing Service (WPS), geographical web services, service chain, composition, SDI, regional planning, housing market analyses.

INTRODUCTION

In the last years the German housing market has become more and more complex. Insufficient knowledge about ongoing developments in this market causes wrong decisions in the housing politics and housing industry (Helsper, 2005). For providing assistance with spatial decisions in this domain, decision makers are supported by the use of *Geographic Information Systems* (GIS). GIS are used to prepare, manage, visualise and analyse relevant spatial and statistical data. Hence, GIS functionalities provide assistance for governmental, economic and administrative decision makers with their assessment and evaluation of the regional and local housing market.

The objective of the project "A Web *Spatial Decision Support System* (SDSS) for the automation of the multi criteria model development of a user specific, regionalised housing market analyses in RLP" is to develop components and tools for a Web SDSS which automates the multi criteria model building for user specific, regionalised housing market analyses in the German State of Rhineland Palatinate (RLP). This system needs to be able to manage and aggregate the necessary spatial and statistical information which is needed to develop supply and demand indicators. The aim of the project is the low budget development of relevant modules for a SDSS for analysing the housing and real estate market in RLP. This is achieved by using and extending open source software packages and the development of own tools. These tools should be transferable to and applicable on other projects. Furthermore, relevant components should be re-usable and conform to the *Open GeoSpatial Consortium* (OGC) and the *World Wide Web Consortium* (W3C) standards.

In this article we want to demonstrate how the OGC standard of a *Web Processing Service* (WPS) interface is applied for spatial analyses in the housing market domain and we want to reveal how flexible GIS functionalities can be used while providing them as WPS processes.

WEB-BASED GIS ANALYSIS WITH OPEN STANDARDS

However, first of all we take a look at the present development in the area of GIS. Traditionally, GIS are monolithic complex systems which are handled by professionals. In recent years advances in information technology have made it possible to move towards *Service Oriented Architectures* (SOAs) and distributed computing for the dissemination of geographic information (Friis-Christensen, 2006). In a SOA environment, resources on a network are made available as independent services that can be used in diverse applications. This interoperability requires the definition of worldwide service standards.

The most important organisation that defines a range of standards in the geospatial domain is the OGC. Amongst others the OGC specified the well-known (partly draft) standards *Web Map Service* (WMS), *Web Feature Service* (WFS), *Geography Markup Language* (GML), *Styled Layer Descriptor* (SLD), *Web Coverage Service* (WCS), *Catalogue Service for the Web* (CSW) and *Web Coordinate Transformation Service* (WCTS). This means that visualisation, vector and raster data access, along with the ability to search for spatial data is mostly covered by these standards. For the processing of this data a standardised service was missing for a long time and the user had to fall back to the usage of his traditional Desktop-GIS. For this reason an OGC working group has developed a further standard which has been recently adopted: the *Web Processing Service* (WPS) interface specification (OGC, 2007a).

The idea behind a WPS is to offer any kind of GIS functionalities for the processing of spatial data. It may provide simple calculations (e.g. the calculation of a buffer) as well as complex computations (e.g. the generation of a climate model). According to the specification a WPS is able to handle more than a single process and there are three mandatory operations performed by a WPS, namely *GetCapabilities*, *DescribeProcess* and *Execute*. The response to a *GetCapabilities* request is an XML-document containing metadata of the WPS and all available processes. A detailed process description as well as input and output parameters are provided for every process as response to a *DescribeProcess* request, also in form of a XML-document. The actual execution of the relevant process is carried out when an *Execute* request is sent to the WPS with all necessary parameters such as the name of the process and further ones, most notably the data to be processed. A central feature of the specification is the possibility to pass this data either directly (typically GML encoded) in the *Execute* request or by a data reference. Such a reference refers to a web accessible resource where the WPS can directly retrieve the data to be processed. Thus, it is possible to avoid the sending of huge amounts of data via the network.

As mentioned above a WPS may provide simple calculations as well as very complex computations. Thus, in principle there are no restrictions on what types of operations can be realised based on the WPS interface. This openness of the specification led to a variety of criticism and discussions in the recent past (e.g. Heier, 2005; Kiehle, 2006) and several research groups work on open issues. In general particular attention is paid at the moment to semantic descriptions of geo-processes and models (e.g. Lemmens, 2006; Diaz, 2008).

GIS FUNCTIONALITY IN THE HOUSING MARKET DOMAIN

In this chapter we want to give two typical examples for analyses in the housing market domain which rely on GIS functionalities, and which we have identified as relevant for our project on the one hand but also for similar typical GIS analyses in possible different domains on the other hand. These examples have been implemented successfully as explained later in section 4. They represent relevant parts of typical analysis workflows in the domain of housing market and consist of a set of more fundamental geodata processing operations and calculations. These offer a toolbox of basic building blocks that can be combined to more sophisticated analysis tasks as explained in the following sections:

Site selection

We start with an example which is well-known by everyone who was ever hunting a house: Site selection. This analysis is important in the housing market domain but it is also a very typical example for many GIS projects and therefore can easily be generalised for other scenarios. In general people are looking for a house which keeps conditions concerning certain distances. A family may look for house near a kindergarten, an elementary school and a supermarket. A man of business may prefer a house near the train station and not too far from the next airport. A typical question could be: “In which area do I have to look for a house which is within a distance of 1.5 km of a kindergarten, within a distance of 2 km of an elementary school and within a distance of 0.5 km of a supermarket?” This kind of question is easily answerable with a GIS based on the respective data. In a first step all kindergartens, elementary schools and supermarkets are buffered by the supposed distance. In the second step all calculated buffers are intersected among each other. These steps are illustrated in figure 1.

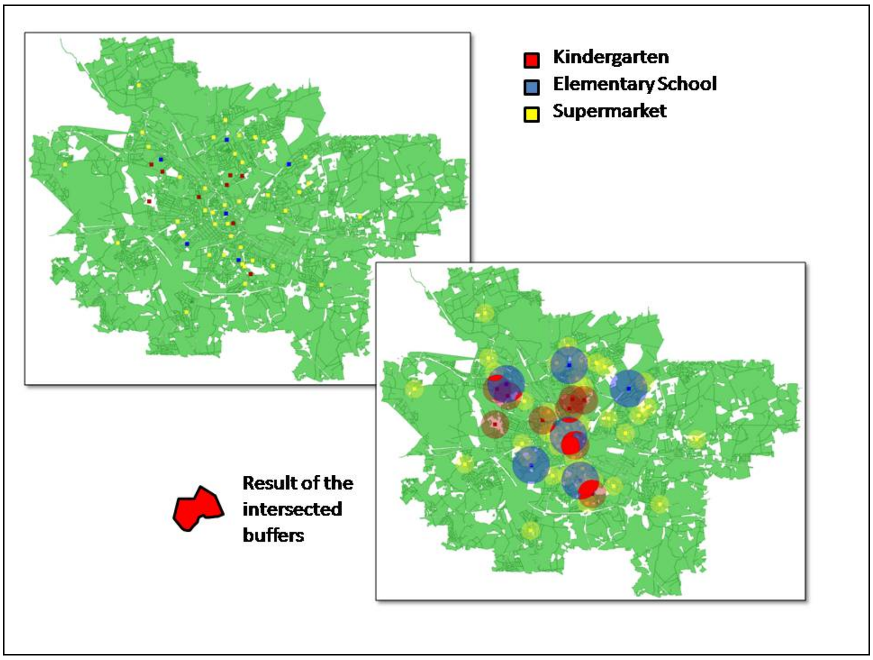

Figure 1: Illustration of the “SiteSelection” procedure.

Granted that there is already a range of potential homes, it is now possible to check whether these objects are situated in the areas calculated before. Through a point-in-polygon query these houses are selected which are finally in the respective distance to a kindergarten, an elementary school and a supermarket specified before. This possible last step is illustrated in figure 2.

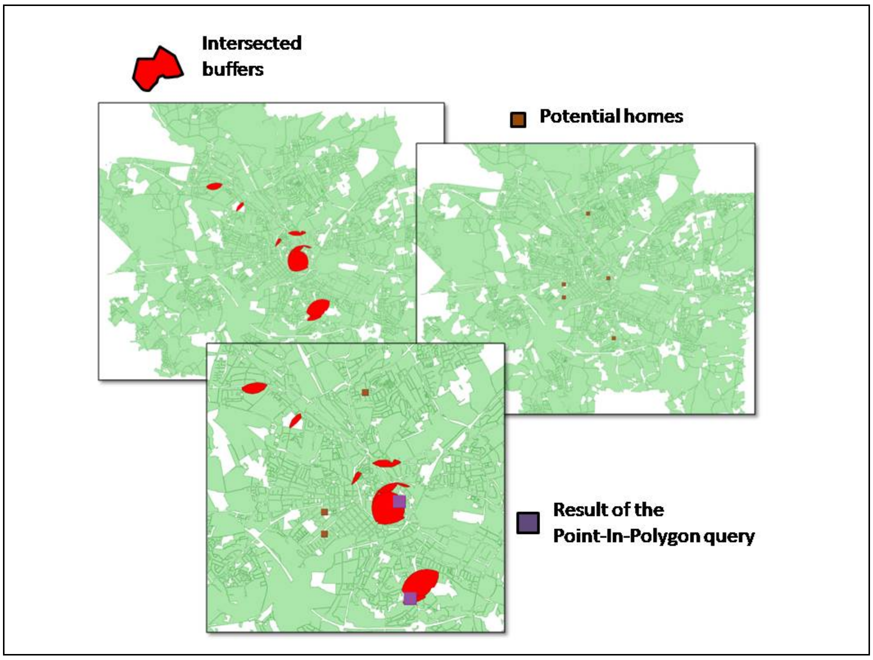

Figure 2: Point-In-Polygon-Query as possible step after the "SiteSelection" procedure.

Regarding the site selection procedure from another point of view, namely off from the private sector and into the economic perspective, we realise that a company has more commercial requirements. For example the location of a new factory must be situated near a highway exit or close to waste incinerator. Given that a low distance to an object does not imply the fast accessibility we have to refine the question put above: In which area do I have to look for a factory site which can be reached within 15 minutes from a highway? Here it becomes clear that a simple buffer operation is not sufficient enough anymore. The accessibility of an object depends on the road network and not on the simple distance. This analysis, namely the calculation of accessibility of regions from a given location, was already realised within our project in a former stage through an *Accessibility Analysis Service* (AAS) (Neis, 2007). By using the AAS instead of a simple buffer operation for a site selection procedure the accuracy of the results is increased, as more realistic driving times on the street network are used there.

Supply areas

In our second example we want to answer another question: Where are demands to be satisfied? This means that specific needs or requirements are identified with the aim to supply these needs. The question is e.g. interesting in the field of urban planning where decision makers analyse in which areas specific demands are not pleased. These demands can for example be kindergarten, supermarkets, doctors etc. and the responsible persons within the urban planning administration have to ensure an adequate supply.

Vice versa we can also regard the question of sufficient supply from the economic perspective: From the view of a product conductor (e.g. a supermarket operator) it is interesting to analyse where to "find" additional customers supposing areas where the demand (of supermarkets) is not satisfied yet. Such an analysis is a typical instrument of "Geomarketing" where GIS tools are used in the process of planning and implementation of marketing activities.

Precondition for the calculation of such supply areas is the availability of population figures for a small scale administrative level because the number of persons is needed to draw further conclusions. In the field of urban planning this data could be building block information which exists for statistical purposes and provides a solid basis for the calculation of supply areas. The whole procedure of this calculation contains two steps: First the providers of interest (e.g. supermarkets) are buffered by a designated distance. In the second step the calculated buffers are intersected with the administrative units (e.g. building blocks). During the intersection the number of residents is aggregated, more precisely the sum of residents is calculated on a pro-rata basis. This means the ratio of the original polygon area and the intersection polygon area are calculated and the value to sum is multiplied by this ratio.

As demonstrated in the site selection procedure above the results are improved by using an Accessibility Analysis Service instead of simple buffer functionality. The procedure is illustrated in figure 3.

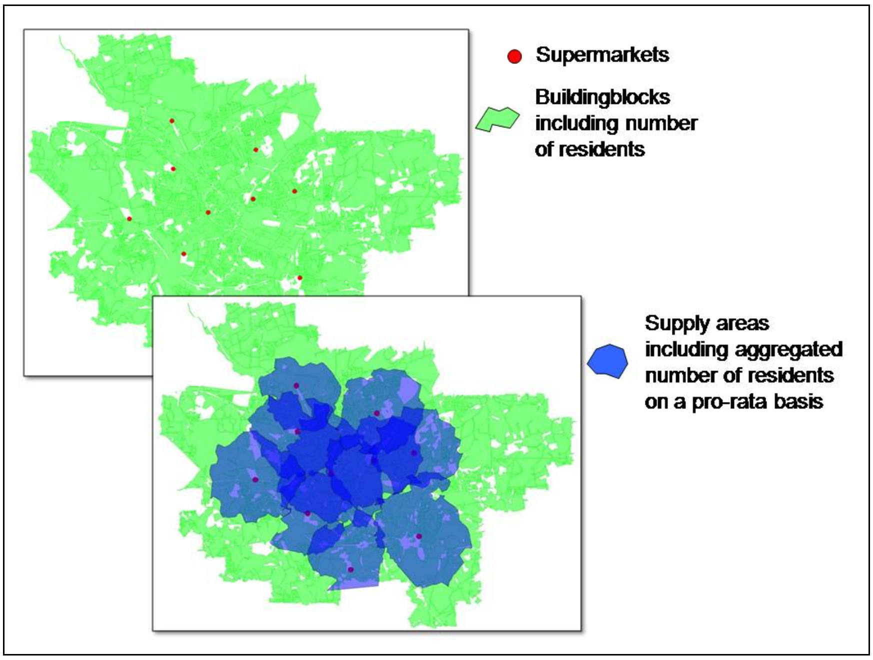

Figure 3: Illustration of the "SupplyAreas" procedure.

USING THE WPS INTERFACE

It became clear that the two described procedures "Site Selection" and "Supply Areas" are partly based on the same fundamental GIS functionalities. One could imagine further applications which are also relying on these basic functions to solve different more complex problems. We used the WPS interface to provide the functionalities mentioned above as independent processes. More detailed we implemented the following basic processes needed for the scenarios which are presented in table 1:

Process	**Description**
Buffer	Creation of a buffer around a GML geometry.
Intersection	Intersection of 2 GML datasets.
PointInPolygonJoin	Join of GML data containing points with GML data containing polygons based on the spatial predicate "contains". All attributes of the points are kept (including the geometry) and the user can choose which attributes of the polygons shall be kept. Furthermore, the user can choose between four different join types: NATURAL, LEFT, RIGHT, FULL. The missing attribute values in cases of left, right or full join are filled with null.
PolygonIntersectsPolygonJoinAggregation	Join of 2 GML datasets containing polygons based on the spatial predicate "intersects" followed by the aggregation of attribute values of the first polygon dataset. Available aggregate functions: count, sum, max and min. In case of the aggregate function "sum" this value is calculated on a pro-rata basis. This means the ratio of the original polygon area and the intersection polygon area is calculated and the value to sum is multiplied by this ratio. It is possible to keep attributes of the second polygon dataset. All polygons of the second dataset (including the geometry) are added to the result data, independent from intersecting polygons of the first dataset or not.

Table 1: Implemented processes for analyses in the housing market domain

These processes are not only used in our current project in the domain of housing market analyses but also in another scenario within a disaster management project. In this so called "BombThreatScenario" we combined some of these basic GIS processes with some services providing routing functionality for the preparation of an evacuation after a bomb finding (Stollberg, 2007). This shows the generic nature of the implemented processes and the applicability within a wide set of different domains.

However, for representing the whole "Site Selection" and "Supply Areas" procedures each as a single application, also further services must be accessed. First of all the relevant data for the procedures must be retrieved, e.g. by accessing adequate WFS. As mentioned in chapter 2 the WPS interface offers the possibility to pass references to web accessible resources providing the relevant data for a process. One possibility for such a reference is the *GetFeature* request to a WFS which is then handled by the WPS for retrieving the appropriate data. Furthermore the Accessibility Analysis Service which defines accessible regions from a location based on calculations on street networks, is accessed. This service is not a standardised *OGC Web Service* (OWS) or even a discussion paper, but its interface is modelled according the OGC OpenLS Implementation Specification. It could also be encapsulated by a WPS, but this is not yet the case in our current realisation.

Service Orchestration

An established standard concerning the *Web Service Orchestration* (WSO) is the use of the *Business Process Execution Language* (BPEL) (Juric, 2004; Chen, 2006). In a former work we discussed already this possibility for the chaining of OWS and encountered some technical problems (Weiser, 2007). One obstacle is the lack of *Web Service Description Language* (WSDL) documents for standardised descriptions of OWS. These WSDL documents act as links between an *Orchestration Engine* (OE) and the service interfaces that are chained. In order to orchestrate an OWS by using BPEL and respectively an OE these WSDL documents have to be created manually for each service involved. Furthermore OWS are so far invoked via HTTP POST and/or GET and not via the *Simple Object Access Protocol* (SOAP). HTTP POST and/or GET are not supported by all Orchestration Engines tested or at least not directly supported. The WPS specification (OGC, 2007a) is one of the first OGC specifications explicitly regarding these problems and describing how to use WPS in connection with WSDL and SOAP. So it will be possible in the near future to orchestrate WPS processes using BPEL and OEs. But in the beginning it will still be necessary to create WSDL documents manually for each process.

In (Stollberg, 2007) we presented an alternative to the usage of BPEL for the orchestration of OWS by using the WPS interface itself. This approach of a "Composite-WPS" we use here again for the combination of all included services (WPS, WFS, AAS) in the "SiteSelection" and "SupplyAreas" procedures.

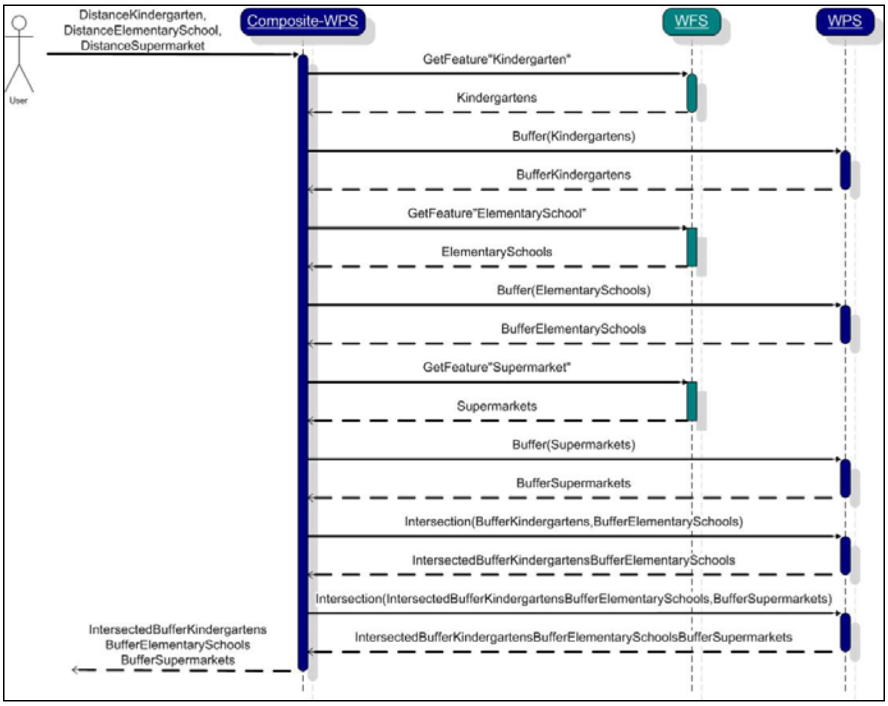

Figure 4: Sequence diagram of the "SiteSelection" process.

The WPS was originally not intended for using it as an application for chaining OWS. But the standard was very open from the beginning and did not exclude any uses of the interface. In the newest version 1.0.0 of the WPS specification it is even explicitly foreseen to use the WPS for the chaining of web services and a WPS can act as a service chaining engine by designing a process that calls a sequence of web services including other WPS processes (OGC, 2007a). We have taken advantage of this in our approach by orchestrating all steps within an independent WPS process. These steps are illustrated for the "SiteSelection" Process in figure 4 and for the "SupplyAreas" process in figure 5.

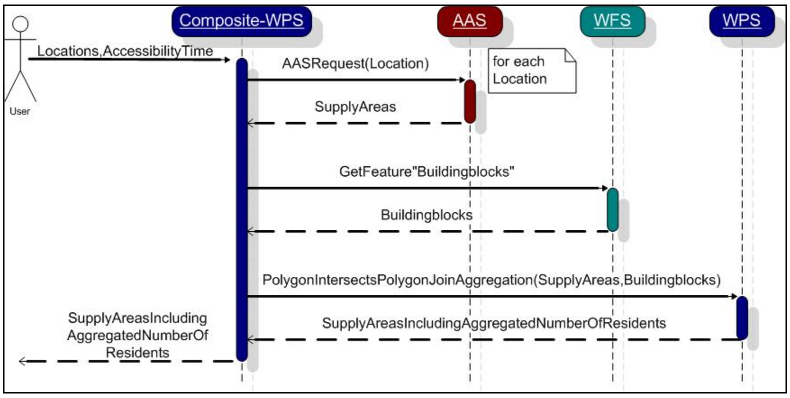

Figure 5: Sequence diagram of the "SupplyAreas" process.

Implementation

For the implementation we used the existing WPS framework of the *deegree2* project (www.deegree.org) (Fitzke, 2004). This and implementations within other projects such as the *52°North* project (http://52north.org) and the *Agriculture and Agri-Food Canada* project (www.agr.gc.ca) are in accordance with the version 0.4.0 of the WPS specification (OGC, 2005).

The *deegree2* framework offers the possibility to implement the application logic of new processes within a separate class derived from the *org.deegree.ogcwebservices.wps.execute.Process.java* class. In addition an XML configuration document has to be created for the process and integrated into the framework along with the process class. After these two steps the new process is then accessible via the WPS interface.

We started with the implementation of the identified fundamental GIS functionalities which are required for our specific "Site Selection" and "Supply Areas" procedures. Afterwards we chained these processes along with requests to WFS and AAS within our "Composite-WPS". The advantage of this approach is that we do not need a BPEL Orchestration Engine and respectively WSDL documents describing the services involved. Both identified domain specific scenarios are implemented using solely OGC specifications. The drawback of this approach is that a lot of work has to be done manually while products like *ActiveBPEL* or *Oracle BPEL* offer a whole toolbox for creating, deploying and managing BPEL business processes including a graphical user interface for the visual chaining of processes.

CONCLUSIONS AND FUTURE WORK

The aim of our project is the development of relevant modules for a SDSS for analysing the housing and real estate market in RLP by using and extending open source software packages and the development of own tools. For this reason we identified a range of GIS basic functionalities which are needed in different procedures of housing market analyses but which are also used within scenarios in different domains. Our modular approach makes it possible to combine single functionalities for the configuration of a flexible Spatial Decision Support System. For this purpose we used the new OGC specification of a Web Processing Service which offers nowadays the possibility for representing complex scenarios completely in accordance with OGC standards. The specification is a big step towards "real" Web-GIS which also includes the actual processing (which means "alteration") of spatial data in contrast to only manage, search and visualise it. Unfortunately the standard is in the current version a bit too open which has lead to complaints in the community. This means that at the moment it is possible to "hide" literally any kind of functionality behind a WPS facade, regardless of being "geo"processing or not. This problem is tackled with the upcoming development of so called "WPS Profiles" which have to be identified for a) basic GIS functionalities and b) domain specific scenarios. The definition of such profiles shall help to ensure that users and providers of WPS processes agree upon specific operations and that the process "Buffer" of Organisation A provides the same functionality and takes the same input and output parameters as the process "Buffer" of Organisation B, thus leading to interoperability. We identified two domain specific operations in housing market analyses which could be proposed for a future possible profile for this specific domain. We could demonstrate that the WPS interface today can be applied successfully in this domain and realised a range of relevant processes. Several further basic processes are currently under development which can be used in this or other domains.

The new OGC WPS standard makes it now possible to provide geoprocessing via a standardised interface, e.g. we are currently working on visibility analysis and the processing and analysis of 3D data which will be offered as WPS processes. There will be in general a quite variety of processes for the WPS user in the near future because the number of WPS processes will increase. For this reason the vision of Web-GIS is approaching. These days it is already possible to find a lot of geographical data from distributed sources via the Web, mostly in the form of maps. In the future we will also find GIS functionality for processing this data according to our requirements. Until now we are obliged to fall back to use our traditional Desktop-GIS for these analyses. But for using the whole potential of geoprocessing functionalities and the flexible combination of granular modules it is still a long way to go because the interoperability is still missing. However, the development of WPS Profiles and respectively standardised geoprocessing operations in the near future will bring us a step further into the direction of interoperable Geo Web Services.

REFERENCES

Chen, L., Wassermann, B., Emmerich, W. and Foster, H. (2006): Web Service Orchestrations with BPEL (Tutorial Summary). In: 28th International Conference on Software Engineering (ICSE 2006), Shanghai, China, May 20-28, 2006, pp. 1071-1072. ACM Press.

Díaz, L., Granell, C. and Gould, M. (2008 in press): Case Study: Geospatial processing services for web-based hydrological applications. In: J.T. Sample, K. Shaw, S. Tu, M. Abdelguerfi (Eds.): Geospatial Services and Applications for the Internet. Springer, 2008. (in press).

Fitzke, J., Greve, K., Müller, M. and Poth, M. (2004): Building SDIs with Free Software - the deegree Project. In: Proceedings of GSDI- 7, Bangalore, India.

Friis-Christensen, A., Bernard, L., Kanellopoulos, I., Nogueras-Iso, J., Peedell, S., Schade, S. and Thorne, C. (2006): Building Service Oriented Applications on top of a Spatial Data Infrastructure – A Forest Fire Assessment Example. AGILE - Shaping the future of Geographic Information Science in Europe, Visegrád, Hungary.

Heier,C. and Kiehle, C. (2005): Geodatenverarbeitung im Internet - der OGC Web Processing Service. GIS 6/2005, ABC Verlag, Heidelberg, pp. 39-43.

Helsper, R., Kehl, A., Mayer, T., Röss, W. und Zipf, A. (2005): Ein GIS-basiertes Web-Portal für die regionalisierte Wohnungsmarktbeobachtung in Rheinland-Pfalz als Basis für ein Spatial Decision Support System (SDSS). AGIT 2005. Symposium für Angewandte Geographische Informationstechnologie. Salzburg.

Juric, M., Mathew, B. and Sarang, P. (2004). Business Process Execution Language for Web Services. Packt Publishing Ltd., Birmingham, 2004.

Kiehle, C., Greve, K. and Heier, C. (2006). Standardized Geoprocessing – Taking Spatial Data Infrastructures one step further. Proceedings of the 9th AGILE International Conference on Geographic Information Science. Visegrád, Hungary.

Lemmens, R., Granell, C., Wytzisk, A., de By R., Gould, M. and van Oosterom, P. (2006): Semantic and syntactic service descriptions at work in geo-service chaining. Proceedings of the 9th AGILE International Conference on Geographic Information Science. Visegrád, Hungary.

Neis, P. and Zipf, A. (2007): A Web Accessibility Analysis Service based on the OpenLS Route Service. In: AGILE 2007. International Conference on Geographic Information Science of the Association of Geograpic Information Laboratories for Europe (AGILE), Aalborg, Denmark.

OGC (2005): Web Processing Service. OGC Discussion Paper, Document Reference Number 05-007r4, Version 0.4.0.

OGC (2007a): Web Processing Service. OpenGIS Standard, Document Reference Number 05-007r7, Version 1.0.0.

OGC (2007b): OWS-4 Workflow IPR. Hrsg. OGC. RefNum OGC 06-187; Version 1.0.0; 2008-04-02 Status: internal OGC Discussion Paper.

Stollberg, B. and Zipf, A. (2007): OGC Web Processing Service Interface for Web Service Orchestration - Aggregating Geo-processing Services in a Bomb Threat Scenario. In: Web and Wireless Geographical Information Systems. 7th International Symposium, W2GIS 2007, Cardiff, UK, November 2007 Proceedings. LNSC 4857, pp. 239-251. Springer-Verlag, Heidelberg.

Weiser, A. and Zipf, A. (2007): Web Service Orchestration of OGC Web Services for Disaster Management. Joined CIG/ISPRS Conference on Geomatics for Disaster and Risk Management, May 2007. Toronto, Kanada.