

## **mdWFS: A Concept of Web-enabling Semantic Transformation**

Andreas DONAUBAUER, Florian STRAUB, Matthäus SCHILCHER,

Fachgebiet Geoinformationssysteme  
Technische Universität München

{ andreas.donaubauer | florian.straub | matthaeus.schilcher } @bv.tum.de

**KEY WORDS:** *Model Driven Approach, OGC Web Services, Schema Transformation, Spatial Data Infrastructures, Service Oriented Architecture, Semantic Interoperability, Conceptual Schema*

### **SUMMARY**

The paper addresses current research issues in the field of interoperability of heterogeneous GI systems. Special emphasis is placed on heterogeneity at the level of conceptual data models – a problem that becomes particularly evident in a cross-border context. A concept of a Model Driven Approach for accessing distributed, heterogeneous spatial data using web services currently being developed in a joint research of the Eidgenössische Technische Hochschule (ETH) Zürich, the Technische Universität München (TUM), the Bundesamt für Kartographie und Geodäsie (German Federal Agency for Cartography and Geodesy) and the Swisstopo (Swiss Federal Office of Topography) is presented. This approach is based on the fundamentals of Service Oriented Architecture (SOA), the Model Driven Approach (MDA) and semantic transformation. This paper concentrates on the design of a web service with the ability to perform semantic transformation at the conceptual schema level. The service interface is discussed by means of a use case.

### **INTRODUCTION**

Spatial Data Infrastructures (SDI) on regional, national and international levels are currently one of the top items on the GI agenda. Fast, easy access as well as efficient and sustainable usage and combination of distributed, heterogeneous spatial information across the borders of vendor systems, disciplines, sectors and countries have become more important than ever and are also stipulated in the directive for an Infrastructure for Spatial Information in Europe (INSPIRE). To achieve these goals, a number of obstacles have to be overcome, one of the biggest being the data's heterogeneity at the level of data models and semantics. This problem is currently not sufficiently addressed by GI standardisation and becomes particularly evident in cross-border web-based GIS applications, where heterogeneous data from sources in different countries need to be harmonised so that they can be combined seamlessly and consistently.

Data harmonisation is also a key aspect of INSPIRE, but “INSPIRE does not require Member States to change the format of their spatial data holdings; instead, Member States can provide interfaces that transform heterogeneous data to a uniform model” (COM 2004). Thus the focus is not on full harmonisation of the underlying data models, but rather on achieving interoperability in a service-oriented architecture (SOA), e.g. by using web services for translating between different data models (Riecken et al., 2003).

In existing web services for using spatial data, as standardised by the Open Geospatial Consortium (OGC) (e.g. Web Feature Service), the syntactic heterogeneity is dealt with by encapsulating the internal structures of heterogeneous GI-systems using standardised interfaces. The internal structures, i.e. also the data models are hidden from the user. On the one hand, hiding the data model from the user leads to an ad hoc easy-to-use access to spatial information. On the other hand, the problem of semantic heterogeneity, is not solved by existing services. To allow for semantic interoperability by translating between different data models in OGC Web Services (OWS), the latter have to be enhanced by adapting concepts of schema translation / semantic translation based on the Model Driven Approach (MDA). This very issue is addressed by this paper which concentrates on a concept of *web-enabling* semantic transformation. The underlying *formalism* for expressing schema mapping rules is not within the scope of this paper, it has been documented in Gnägi et al., 2006.

## THE FUNDAMENTALS OF SOA, MDA AND SCHEMA TRANSFORMATION

Service Oriented Architectures (SOA) and the Model Driven Approach (MDA) are both common concepts in software development.

The term SOA describes a concept for a system architecture that provides functionality through services in a network, e.g. the web. The components in a SOA are in most cases distributed and heterogeneous and can be coupled loosely, which makes them reusable. Service requests and responses can be stringed together (referred to as “service chaining”) to represent complex processes. The functionality provided by the service can be accessed via standardised interfaces that encapsulate the internal structures of the underlying system and thus facilitates interoperability between the systems (Straub 2005). In the GIS domain, the Open Geospatial Consortium (OGC) specifies such interfaces.

The main principle of the Model Driven Approach (MDA) in software development is designing a precise, platform independent model (PIM) for the software to be developed. The PIM can be used to generate the actual executable software components by automatic transformation (platform specific model, PSM). The main goal of the MDA is to facilitate the design and implementation of platform-independent software components (Straub 2005).

The MDA can also be applied to data interchange. The approaches for transferring data range from the relatively simple format-based data transfer to the powerful model-driven data transfer. In the latter, additional information e.g. on the conceptual model of the data can be transferred together with the data, whereas in the first approach only the transfer format of the data is described. To facilitate a model-driven data exchange, the data’s conceptual models (i.e. information about the structuring in classes and the corresponding attributes) have to be described in an exact, formal and machine readable way using a conceptual schema language (CSL) (e.g. the Unified Modeling Language UML) and rules for automatically deriving a transfer format from the conceptual schema. Processing engines which are able to read and write the transfer format have to be available.

Processes and techniques required to transparently view and query data from multiple, heterogeneous data sources that provide data on the same types of real world objects in different data models as one uniform data source are called semantic transformation.

Semantic transformation approaches can be classified according to the following criteria (Donaubauer et al., 2006):

1. *Level of abstraction*: Semantic transformation can be performed on different levels of abstraction (conceptual, logical, physical level). Semantic transformation on conceptual level is platform independent whereas approaches on logical and physical levels are platform specific.

2. *Orientation*: Semantic transformation can be performed either horizontally (between different schemas on one level of abstraction, PIM to PIM or PSM to PSM) or vertically (between different schemas on different levels of abstraction, PIM to PSM).
3. *Level of automation*: Semantic transformation can be performed on different levels of automation as far as the mapping between entities in different schemas is concerned. Normally the mapping is carried out by hand but there are also some approaches for matching schemas automatically often involving ontologies.

Semantic transformation is identified to be a key interoperability issue in international SDI-related initiatives and projects such as INSPIRE, RISE and EuroSpec. However, no detailed specifications concerning type and level of the transformation or tools for executing them have been made yet. In a number of international projects, e.g. OGC GOS-TP (OGC, 2003a; OGC, 2003b), GiMoDig (Sarjakoski and Lehto, 2004) and SDIGER (Orlova and Bejar, 2005), web services are used to translate horizontally between different schemas (e.g. translating road data from a national schema to a common schema) and prototypes of so-called “translating WFS” have been implemented.

All the transformations implemented in the above mentioned projects were executed on the logical level (e.g. using Xquery or XSLT). Translation between different conceptual schemas (PIM to PIM), the automatic generation of a transfer format (PIM to PSM) and the definition of a language to express these translations were identified as future fields of research. The project mdWFS<sup>1</sup> aims at tackling these fields and facilitating web based transformations on the conceptual level (PIM to PIM and PIM to PSM).

## CONCEPT OF WEB-ENABLING SEMANTIC TRANSFORMATION

Web-enabling semantic transformation in our case means designing a web service. There are two main requirements that have to be fulfilled by the service:

1. ability to provide access to geospatial feature data based on the data’s original conceptual schema (source schema) as well as on any user-defined conceptual schema (target schema)
2. interoperability with existing OGC Web Services

Taking into account these requirements, we designed a service that we call mdWFS (model driven Web Feature Service). The task of this service is to carry out semantic transformations (PIM to PIM mapping) resulting in GML application schemas (PIM to PSM mapping) that will be served by OGC Web Feature Services. The mdWFS has two operations:

1. *GetCapabilities* allows any mdWFS client to retrieve metadata about the capabilities provided by any mdWFS server. Syntax and semantics of this operation are defined in the OGC Web Services Common Specification (OGC, 2005b). Instead of the feature type list provided by the GetCapabilities response of an OGC WFS, the mdWFS GetCapabilities response contains a list of all conceptual schemas provided by the specific mdWFS instance (source schema and target schemas). As each conceptual schemas is represented by a GML application schema and a WFS instance respectively, each item in this list points to a WFS instance. The WFS instance can then be applied to request the corresponding conceptual schema using the DescribeFeatureType operation. The output format of this operation will

---

<sup>1</sup> With the goal of combining concepts of SOA and MDA in the GIS domain the Technische Universität München (TUM) and the Eidgenössische Technische Hochschule Zürich (ETH) have teamed up in the research project „Model driven approach for accessing distributed spatial data using web services - demonstrated for cross-border GIS applications” (mdWFS), funded by the Bundesamt für Kartographie und Geodäsie (German Federal Agency for Cartography and Geodesy) and the Swisstopo (Swiss Federal Office of Topography).

be the XML Metadata Interchange (XMI) (OMG, 2005b) used as a textual representation of the underlying conceptual schema.

2. *TransformSchema* carries out the semantic transformation based on the following input parameters: the source schema identified by the DescribeFeatureType request URL of the WFS instance representing the schema. The target schema encoded in XMI and the schema mapping rules (PIM to PIM) encoded in a formal language based on MOF-QVT (OMG 2005a) documented by Gnägi et al., 2006. The PIM-to-PSM mapping (conceptual schema to GML application schema) is defined by standardised encoding rules. In response to a TransformSchema request, the mdWFS informs the client about the successful completion of the schema transformation and provides the GetCapabilities request URL as a unique identifier of the new WFS instance that corresponds to the conceptual schema resulting from the schema transformation. This WFS instance can then be used to request data by a standard WFS client. The conceptual schema resulting from the schema transformation is published by the mdWFS in its GetCapabilities response (see GetCapabilities operation described above) and can then also be used as a source schema for a new TransformSchema request.

## USE CASE

Figure 1 shows the usage of a mdWFS in the context of a cross-border spatial data infrastructure (SDI).

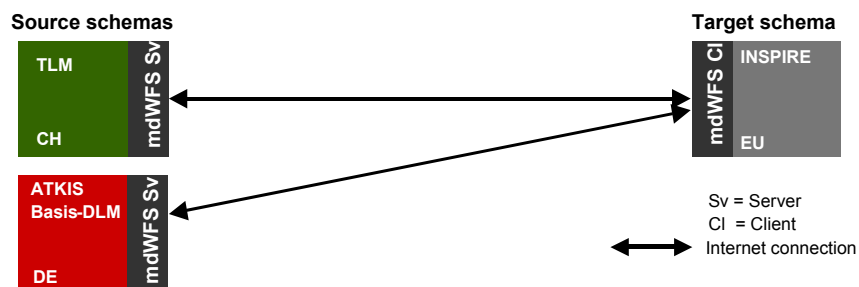


Figure 1: Usage of a mdWFS in the context of a cross-border spatial data infrastructure (SDI)

In our scenario, a user working at a government agency of the European Union needs information on administrative boundaries in the Lake Constance region. In order to get a full coverage of the geographic area of interest, the user has to access two different topographic information systems (“TLM” and “ATKIS Basis-DLM”), maintained by the Swiss and the German mapping authorities respectively. Three different conceptual schemas are involved in this scenario: the original schemas of the Swiss and the German information systems (named source schemas in figure 1) as well as the user’s own schema (named target schema in figure 1).

Provided that the Swiss and the German mapping authorities each operate a mdWFS server (named mdWFS DE and mdWFS CH in figure 2), the workflow of this scenario is as follows:

The user requests the capabilities of both servers and recognizes that the mdWFS DE is already able to serve its data corresponding to the desired target schema, i.e. somebody else has already evoked the TransformSchema request resulting in a new WFS instance (named WFS EU, see left hand side of figure 2). The WFS EU can then be used by a standard WFS client.

In case of the mdWFS CH (right hand side of figure 2), a schema transformation between the Swiss conceptual schema and the EU one has not been carried out yet. Therefore, the user first has to request a description of the Swiss schema (DescribeFeatureType request, OUTPUTFORMAT=XMI). The user then formulates mapping rules between the Swiss schema and the EU schema which are sent

to the mdWFS together with a description of the target schema using the TransformSchema operation. The translation is executed in the mdWFS, using the mapping rules and the description of the target schema. The transformation results in the initialization and configuration of a new WFS instance (in figure 2 named WFS EU). Thus the data from the Swiss source system can be delivered to the user in the EU agency corresponding to the desired EU target schema.

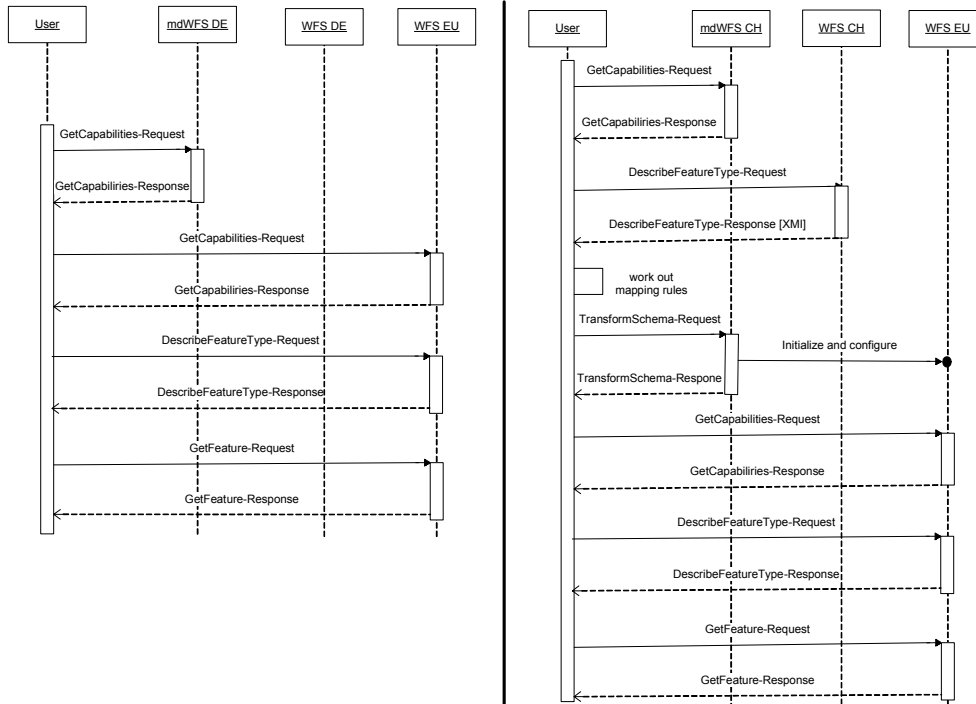


Figure 2: Usage of a mdWFS in the context of a cross-border spatial data infrastructure (SDI)

## CONCLUSIONS AND FUTURE WORK

The paper presents a concept of web-enabling semantic transformation. In contrast to other approaches for achieving semantic interoperability, the semantic transformation is established at the conceptual schema level. The web service defined by the concept is based on existing GI standards like WFS, GML which allows for integration in an OGC Web Service environment. These specifications are extended by MDA-specific standards like XMI and MOF-QVT.

The concept relies on the CSL-descriptions of geographic data sets. Comparing the topographic information systems of Germany and Switzerland, the project mdWFS showed that the current evolution of these systems fulfil this precondition but the comparison also showed that the meta models of the applied conceptual schema languages differ in detail – a fact, the concept must cope with.

Our future work within the project mdWFS will be based upon a prototypical implementation of the concept described in this paper. Our research will cover the following topics: internal system architecture and performance issues of a web-enabled semantic transformation service, further development of the formal language for expressing mapping rules (Gnägi et al., 2006) and evaluation of the prototype in the contexts of cross-border GIS applications and INSPIRE.

## BIBLIOGRAPHY

- COM 2004 = Commission of the European Communities (ed.) 2004, Proposal for a Directive of the European Parliament and of the Council establishing an infrastructure for spatial information in the Community (INSPIRE), COM(2004) 516 final, 2004/0175 (COD), SEC(2004) 980
- Donaubauer, A.; Fichtinger, A.; Schilcher, M.; Straub, S.: Model Driven Approach for Accessing Distributed Spatial Data Using Web Services - Demonstrated for Cross-border GIS Applications (In: Proceedings of the XXIII International FIG Congress, 8 - 13 October 2006)
- Gnägi, H.R.; Morf, A.; Staub, P. 2006, Semantic Interoperability through the Definition of Conceptual Model Transformations, 9th AGILE Conference on Geographic Information Science, 20-22 April 2006, Visegrád
- OGC 2003a = Open Geospatial Consortium Inc. (ed.) 2003a, GOS-TP Translating WFS (WFS-X) - Galdos's Experience.
- OGC 2003b = Open Geospatial Consortium Inc. (ed.) 2003b: WFS-X Schema Translation: The GOS-TP Experience.
- OGC 2005a = Open Geospatial Consortium Inc. (ed.), Web Feature Service Implementation Specification, Version: 1.1.0
- OGC 2005b = Open Geospatial Consortium Inc. (ed.), Web Services Common Specification, Version: 1.0
- OMG 2005a = Object Management Group Inc. (ed.), Meta Object Facility (MOF) 2.0 Query/View/Transformation Specification, November 2005. At <http://www.omg.org/cgi-bin/doc?ptc/2005-11-01>
- OMG 2005b = Object Management Group Inc. (ed.), MOF 2.0/XMI Mapping Specification, v2.1, September 2005. At <http://www.omg.org/docs/formal/05-09-01.pdf>
- OMG 2006 = Object Management Group Inc. (ed.), Meta Object Facility (MOF) Core Specification, Version 2.0, January 2006. At [http://www.omg.org/technology/documents/formal/MOF\\_Core.htm](http://www.omg.org/technology/documents/formal/MOF_Core.htm)
- Orlova, N; Bejar, R. 2005, SDIGER project, INSPIRE Ad hoc meeting, 8 November 2005, Bonn
- Riecken, J.; Bernard, L.; Portele, C.; Remke, A. 2003, North-Rhine Westphalia:
- Riecken, J.; Bernard, L.; Portele, C.; Remke, A. 2003, Building a Regional SDI in a Cross-Border Environment / Ad-Hoc Integration of SDIs: Lessons learnt, 9th EC-GI & GIS Workshop, 25-27 June 2003, A Coruña, Spain. At [http://www.ec-gis.org/Workshops/9ec-gis/papers/pd\\_portele.pdf](http://www.ec-gis.org/Workshops/9ec-gis/papers/pd_portele.pdf)
- Sarjakoski, T.; Lehto, L. 2004, The GiMoDig Schema translation based on XSLT, EuroSDR Workshop, 16 April 2004, Paris. At <http://www.laserscan.com/eurocdr/ontologies/pres/C1-GiMoDig-schema-trans.pdf>
- Straub, F. 2005, Modellbasierter Ansatz für den Web-Zugriff auf verteilte, objektstrukturierte Geodaten, Master's Thesis, Technische Universität München, Fakultät für Informatik, Munich