WISDOM DEMONSTRATION MODULE – OPEN SOURCE TECHNOLOGY MANAGING WATER RELATED, SPATIAL INFORMATION

Thilo Wehrmann[1], Tran Thai Binh[3], Steffen Gebhardt[2], Juliane Huth[2], Claudia Künzer[2], Lam Dao Nguyen[3], Ingo Schettler[1], Michael Schmidt[1] and Stefan Dech[1]

[1]German Remote Sensing Data Center (DFD), German Aerospace Center (DLR)
Münchner Strasse 20, 82234 Weßling, Germany
Corresponding email: thilo.wehrmann@dlr.de
[2]Chair of Remote Sensing, Institute of Geography,
Julius-Maximilians-Universität Würzburg,
Am Hubland, 97082 Würzburg, Germany
[3]GIS & Remote Sensing Research Center (GIRS), HCMC Institute of Resources Geography (HCMIRG),
Vietnamese Academy of Science and Technology (VAST)
1 Mac Dinh Chi St., Dist. 1, HCMC, Vietnam

ABSTRACT

Nowadays open source technologies and applications in the field of Geoinformatics permit the development of spatial data infrastructures (SDI) and internet applications even for comprehensive data sources and complex business processes, especially in the context of projects with developing countries. In the context of the WISDOM project – a German-Vietnamese bilateral initiative – a prototypical spatial data enabled application is designed and implemented using free software components. The demonstration module can be seen as a concept study that targets Vietnamese stake holders in the water related sector. The utilisation of easy-to-use web applications in combination with up-to-date Earth Observation data, hydraulic models and results from socioeconomic surveys supports administrational activities in the background of the integrated water resource management in the Mekong Delta.

1. INTRODUCTION

Spatial data and geo information are used in environmental planning, utility companies, planning offices, and to support spatial decision making (Kiehle, 2006). Recent technological developments have made it possible to create a genuinely open architecture through the use of standard components (Dunfey et al, 2006). Beside monetary aspects the main potential of open source software lies in accessing source code directly for own modifications and further developments, optimisation of processes based on user requirements and steady development of software due to large international communities. The WISDOM project (Water related information system for the Mekong Delta) is an bilateral project between Germany and Viet Nam (http://www.wisdom.caf.dlr.de). The aim of the project is the design, implementation and training of technologies in the integrated water and resources management domain. One aspect is a prototypical development of a water related information system for governmental institutions in the Mekong Delta. The information system consists of harmonised spatial data management, operational processing of remotely sensed data, contextual data models including socio economic data (Semantic web) and user friendly graphical interfaces via web browses (thin client) or rich web application (thick clients). The paper will address the system architecture used for the demonstration module, the experiences using OGC compliant services we gained during implementation and the exemplified operation of a water related information system in Vietnamese governmental institutions.

2. SYSTEM ARCHITECTURE

In software engineering one important aspect is encapsulation of the data model, the application logic and the interface to the user (Hofmann, 1999). Beside Model-view-controller (MVC) concept this can be achieved using multi-tier architecture, also known as client-server architecture. The use of this type of architecture, resp. design pattern permits the modular implementation of complex applications, including distributed web applications. The typical scenario is the 3-Tier architecture, including Data Tier, Application Logic Tier and the Presentation Tier.

2.1 Data Tier

The so called back end in the system architecture handles data storage and data management of the application. Often data base management systems are used for that purpose. Using Structural Query Language (SQL) or Object Query Language (OQL) derivates relational structured or object-oriented data sets can be queried, filtered and altered to the users` needs. Beside data itself information about the data or metadata can be stored in this tier. Metadata is necessary for structuring the content of the data and their relationship to other data sets which cannot be represented in the data model itself or the implementing table design. Moreover application related data can be also stored in this tier to provide persistency, respectively hibernation and enables user right management, to mention just a few examples.

In most cases business objects and the relating metadata can be stored in relational table structures. Even spatial data (data with geographic locations) are managed in modern relational database management systems like the spatial extensions of Oracle and Microsoft and the free PostGIS module.

2.2 Application Logic Tier

All necessary operations and actions of the application are controlled in the application layer or middle ware. It is the translator between the users´ needs and the corresponding data. Beside the business logic, interfaces are necessary for inter tier / inter process communication. The communication of web services can be based on adaptations of the Extensible Markup Language (XML) for data description like the Simple Object Access Protocol (SOAP) or Representational State Transfer (REST) architectures for controlling, resp. exchange between different processes. Business logic consists of business objects and business processes. To simplify the design of complex processes in SOAP environments the Web Services Description Language (WSDL) and WSDL engines describe the resources and support process orchestration and execution. One common use case for the proposed information system is a general status report generation for different water-related topics. For it maps are created with dynamic layer content, statistics are calculated based on recent Earth Observation (EO) data or census data and task conform documents are compiled for report distribution.

2.3 Presentation Tier

The presentation layer or front end of the application is the actual interface to the end user. All specific actions, like initiating and receiving requests are performed in this tier. The prepared content is visualised in the client application. Typical web applications utilise well developed components of the Internet, e.g. Hypertext Transfer Protocol (HTTP) and Uniform Resource Identifiers (URI). The main advantage is the integration of comprehensive techniques and existing infrastructures for the development of custom applications. There are two types for web-based client applications: web browser based thin clients and Rich Internet Applications (RIA). Applications of the first type don’t need special software installations beside a modern web browser. Complex applications with interaction to the operation system need ‘thick’ clients.

2.4 Spatial Data Infrastructure

The concept of Spatial Data Infrastructure (SDI) describes a common framework how spatial data, its metadata and the corresponding processes are interlinked with each other in a scalable and distributed. There exists a set of large scale SDI initiatives way (Masser, 1999), among the National Spatial Data Infrastructure (NSDI) in the US and Infrastructure for Spatial Information in Europe (INSPIRE) for the European membership countries. Further information can be accessed from the Global Spatial Data Infrastructure Association (GSDI). The connection of different data sources from different data providers involves common standards for data exchange, data description, and process communication, to name just a few examples. These standards are in focus of International Standardisation Organisation (ISO) or the Open GIS Consortium (OGC), a non-profit initiative of leading spatial application manufactures. Both boards collect requirements and define standards for various topics in the field of Geo Information. OGC specifies a set of services – common process interfaces – for different applications (direct data access, map visualisation, process controlling, etc.) Using these standards it is possible to couple

3. WISDOM DEMONSTRATION MODULE

In the WISDOM project exemplarily water relevant data and information from the Vietnamese partners and participating organisations are collected, pre-processed and indexed. After data harmonisation, data is stored in distributed database systems for every organisation. For the data catalogue system metadata is automatically generated and stored in the ISO 19115 standard. In parallel hydraulic models are calculated for the Mekong Delta area calibrated with buoys measurements and remote sensing data. Beside water extend and quality additional data products are generated based on remote sensing data, including land cover classification. In the WISDOM project a demonstration module (available at http://demonstrator.wisdom-project.org) was build as a technical case study using this three tier architecture and OGC services for data exchange. Figure 1 shows the general structure with the elements (a) topic access, (b) additional information products and (c) map / data view. All existing data sets are stored systematically in database structures or in directory structures, including remote sensing data, remote sensing products (land cover classification, water covered areas and water quality), hydraulic model results (1D and 2D) and socio-economic data (inter alia population distribution and legal frameworks).

The demonstration system itself consists of three virtualised systems using free VMware Server 1.0.5 and Linux Ubuntu 7.04 as operation system: database VM, Web server VM, and Mapserver VM. The advantages of splitting these components are load sharing, scalability and, pragmatically, application installation. The current installation avoids using a virtualised database system due to better performance. File storage for raster data is distributed via common network protocols (NFS and CIFS).

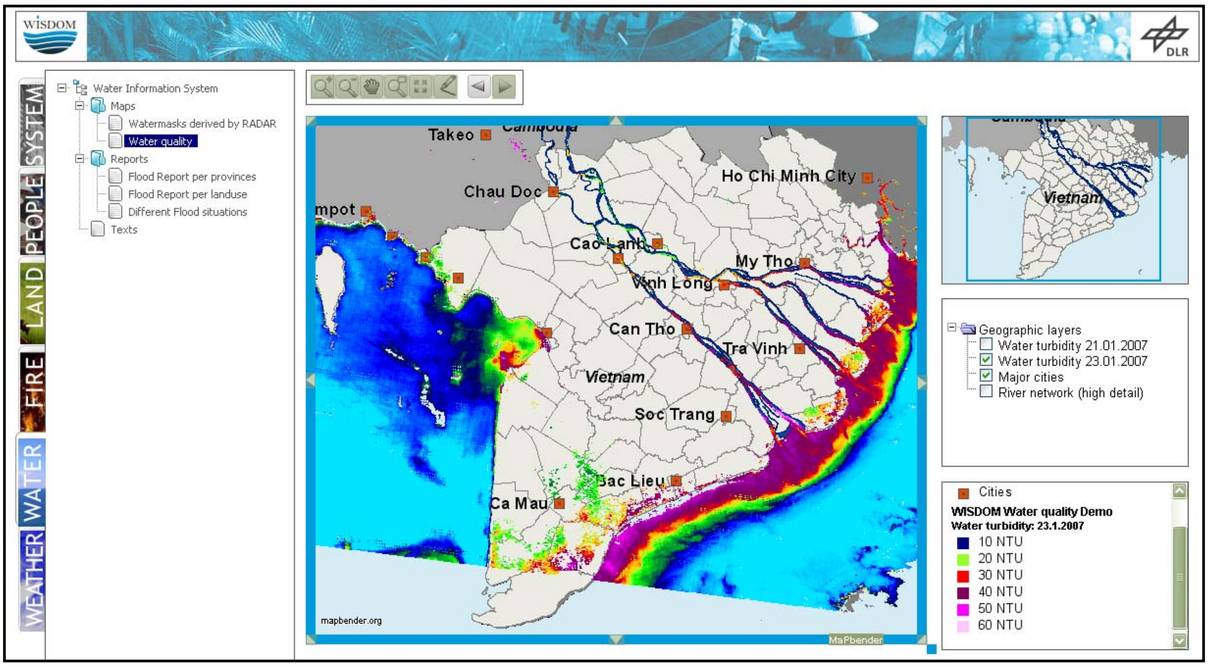

Figure 1. Web browser GUI of the demonstration module

3.1 Data management

PostgreSQL (http://www.postgresql.org) is an open source, object-relational database management system supporting SQL92, transactions and a procedural language. It has build-in geometric primitives like point, line segment or polygon and with PostGIS extension (http://postgis.refractions.net), a spatial database extension, also many GIS functionalities (e.g. intersection, union, spatial queries, etc.) The demonstration module uses PostgreSQL v8.1.11 and PostGIS v1.1.6 with GEOS and Proj library. The integration of available and newly generated geospatial data into the information system appeared to be difficult because of heterogeneities in data formats, map projections, and metadata description. In SDIs, existing standards often fail to address semantic problems that occur due to heterogeneous data content and heterogeneous user communities, using different languages, terminologies, and perspectives (Lutz et al., 2008). Geodata must be current, of known quality, easily accessible, and able to be integrated, generalized, and manipulated with known or predictable results and therefore require some degree of standardization of data models, data content, feature delineation, data collection, georeferencing, indirect positioning, data quality, metadata, and data transfer and exchange (Guptill, 1994) All collected data in the project was pre-processed according to a project-wide data standardisation effort. The existing spatial data formats and projections were restricted to GeoTIFF, respectively ESRI shape file and Geographical projection / UTM zone 48 (both with WGS84 ellipsoid and datum) based on the WISDOM Geodata Exchange Format (WGEF). For every dataset, an associated metadata entry was created, too. The processed vector data was dumped into the PostgreSQL database into two schemes: *wdm_geodata.public* for the static content and *wdm_geodata.processed* for updated data (e.g. fire points) and newly generated data (spatial data analysis).

The UMN Mapserver (http://mapserver.gis.umn.edu) is an OGC compliant mapping server providing OGC services for spatially enabled internet applications. The Mapserver application was originally developed by the University of Minnesota, USA and is now part of the OSGeo project (Open Source Geospatial Foundation). The server supports the common OGC services, including Web Mapping Service (WMS), Web Feature Service (WFS), and Web Coverage Service (WCS). For the demonstration module only the mapping service with its advanced cartographic capabilities using Styled Layer Descriptor (SLD) language was used (see Lee, 2005). The GeoNetwork application (http://geonetwork-opensource.org) is a comprehensive catalogue system for the management of spatially referenced data resources based on OGC conform interfaces and also a part of OSGeo. The GeoNetwork software was developed by UN organisation FAO in the year 2001. The purpose was metadata search, administration, data distribution and publication and an interactive access to the data using standard technologies and data formats. For the demonstration module GeoNetwork v2.0.3 was installed.

3.2 Middleware

The dynamic generation of maps and reports requires a uniform data access to the existing web services, respectively the database system / file system of the data tier. The implementation of the business logic was realised with PHP, a script-based programming language for web applications. The assumed business case for the demonstration module matches administrational report production based on use cases like flooding scenarios planning. Static data as XML schemes, graphics and so on, is distributed via apache http server. As web server it also handles CGI access to the Mapserver component.

As for the statistical reporting PHP as server-side html embedded scripting language along with JpGraph, a object-oriented graph creating library (http://www.aditus.nu/jpgraph), have been utilised for on-the-fly performing spatial queries on datasets in the spatial database, calculating, and rendering statistical reports including map overviews retrieved from the WMS services, tables and bar charts. Such report is e.g. presenting a quantitative estimation of flooded areas with respect to the land use type. Using dynamic processing it is ensured that recent data sets are used for report generation.

3.3 Web client

The aim of developing a web based client was the prevention of software installation in combination with an easy to use application. The DHTMLX toolkit (http://www.dhtmlx.com) consists of a wide range of user interface components for browser based applications, like toolbars, menus, tree structures, tabs, etc. All these components are JavaScript based and therefore client-side performed using Asynchronous JavaScript and XML (AJAX) techniques, allowing more interactive web applications. The framework enables the end user to focus on certain topics like People, Fire, Water, Land and Weather for exploring the data. Mapbender (http://www.mapbender.org) is an OGC compliant WMS client application based on PHP and JavaScript. It provides a sophisticated data model, management functionality and interfaces for data visualisation, navigation and querying (Christl, 2008). All interactive mapping functionality of the demonstration module is based on the Mapbender v2.4.3 client.

4 OUTLOOK

Besides gaining experience building spatial data enabled web applications the intention of the demonstration module was to indicate German and Vietnamese project partners the abilities of a web based application as presentation layer for their data and results. For Vietnamese stake holder the presentations should point to the potentials of combining data from different topics for solving recurring duties. In both points the application could convince the audience and support the understanding the requirement of water-related information as well as technical issues.

After the successful demonstration the first prototype of the proposed information system was designed based on the achieved experiences. The focus of the upcoming prototype is the integration of non-spatial data and the dynamic processing and visualisation of periodical updating data sources. Therefore the prototype consists of the following aspects: ingestion of constant data flow into the system, operational data processing of dynamic data sets, and interactive data exploration of various topics.

5 CONCLUSION

With existing frameworks by OGC and comprehensive libraries it is possible to implement spatial enabled web based applications for research projects like the WISDOM project. The necessary effort was done in three months and the upcoming work load was distributed to four people for data preparation, module design, implementation and testing. The vector data processing in the database tier show insufficient performance for the report generation. Parts of the processing had been statically processed. The generation of Near-Real-time (NRT) data within this framework has to be analysed in the following prototype phase. Moreover several tiny, more technical hitches have pointed out like browser incompatibilities and the adoption of existing software like Mapbender and GeoNetwork.

6 ACKNOWLEGEMENTS

The WISDOM project is funded by the German Federal Ministry of Education and Research (BMBF) (FKZ 0330777) and the Vietnamese Ministry of Science and Technology (MOST).

7. REFERENCES

Christl, A., 2008, Introduction to Geoportal Management using Mapbender, *FOSS4G Proceedings* (in press)

Dunfey, R., Bruce M. Gittings, James K. Batcheller, 2006. Towards an open architecture for vector GIS. *Computers & Geosciences* 32(10), p. 1720-1732

Guptill, S.C., Spatial data standards and information policy. Government Information Quarterly, 1994. 11(4): p. 387-401.

Hofmann, C. 1999, A multi-tier framework for accessing distributed, heterogeneous spatial data in a federation based EIS. *Geographic Information Systems*. Proceedings of the 7th ACM international symposium on Advances in geographic information systems.

Kiehle, C., (2006) Business logic for geoprocessing of distributed geodata. *Computers & Geosciences* 32 (10), p. 1746–1757

Lee, H., and Hyunah, L. 2005, Design and implementation of integrated MapServer in Web mapping environments. *Geoscience and Remote Sensing Symposium*, 2005. IGARSS 05, Proceedings

Lutz, M.J. Sprado, E. Klien, C. Schubert, I. Christ (2008). Overcoming semantic heterogeneity in spatial data infrastructures. *Computers & Geosciences* (in press)

Masser, I. (1999). All shapes and sizes: the first generation of National Spatial Data Infrastructures, *International Journal of Geographical Information Science*, 13:67-84.