

Finding geodata that otherwise would have been forgotten – GeoXchange – a portal for free geodata

Sven Tschirner and Alexander Zipf

University of Applied Sciences FH Mainz
Department of Geoinformatics and Surveying
i3mainz - Institute for Spatial Information and Surveying Technology
Holzstrasse 36
55116 Mainz, Germany
zipf@geoinform.fh-mainz.de
<http://www.geoinform.fh-mainz.de/~zipf/>
www.geoXchange.org

Abstract. A key problem of small organizations like departments of geography, ecology, geodesy, architecture, landscape planning, archeology etc is, that they also produce – from time to time – some spatial data – mostly in small projects or master theses or PhD thesis. This data is usually not included in any spatial data infrastructure. Therefore we developed a GeoPortal – the *geoXchange* Portal - that gives the possibility to store, annotate and share such geodata produced by small institutions that do not have the resources to build up a SDI on their own. The portal is realized using open standards of OGC, ISO and W3C. It allows describing the geodata sets with metadata according to ISO 19115 (CORE) – as well as Dublin Core. This metadata can be queried using a Web Interface to an OGC conformant Catalog Service (CSW) and also using a graphical map interface (bounding-box) through a Web Map Service (WMS). Geodata itself can be uploaded to the portal directly through a transactional Web Feature Service (WFS) through which it also can be accessed. Uploaded data is made automatically available via the map query interface, too, in order to allow visual queries.

1 Introduction

A lot of initiatives, organizations and projects start to develop spatial data infrastructures (SDI) in order to ease access and use of their spatial data and geo-services. Most of these projects care for geodata that is being generated either by federal agencies or by big organizations. A key problem of small organizations like departments of geography, ecology, geodesy, architecture, landscape planning, archeology etc is, that they also produce – from time to time – some spatial data – mostly in small projects or master theses or PhD thesis. This data is usually not included in any spatial data infrastructure. In the contrary, in nearly all time the data is forgotten and lost after the respective project or thesis finishes or the person directly involved with generating the data has left the organization. While the individual dataset that has been produced

may be small, the cumulative loss of all that data sets over the years on an international scale is enormous and not acceptable. The generated data sets usually are not even put to the Web and even when they usually are not found by search engines (except for exceptions like GeoMeta.Info (xxx)). At least they are not found using standard catalog services using structured search capabilities including parameters like: BoundingBox, Thematic and temporal aspects. Further the data comes in a variety of formats, leaving the future user the task to convert this into the desired format and spatial reference system. The only solution is to make these kinds of data sets available worldwide via standardized interfaces and Metadata. While indeed such data sets from research results are sometimes published on the web, the data itself isn't documented at all or even if it is it is badly. This means that it will either be not found at all in the first place – or even when it has been found, it cannot be used adequately because of the lack of metadata which renders the whole dataset useless after the person that produced the data is no longer available. We therefore developed an first version of a GeoPortal – the geoXchange Portal - that gives the possibility to store, annotated and share geodata of these small institutions, projects or persons that usually do not have the resources to build up a spatial data infrastructure (SDI) on their own. The portal is realized with open source software and developed using open standards of the OGC, ISO and W3C. Most important it allows to describe the geodata sets with metadata according to ISO 19115 (currently CORE).

This metadata can be queried using a Web Interface to an OGC conformant Catalog Service (CSW) and also using a graphical map interface through a Web Map Service. The data itself can be uploaded to our Portal directly through a transactional Web Feature Service (WFS) through which it also can be accessed. As the WFS itself only supports GML we additionally realized the automatic Web-based conversion from other proprietary datasets to GML and vice versa in order to ease the use of the portal for people without the technical background or access to a full GIS package. During the upload process the geodata is usually converted to a common Spatial Reference System (CRS) and by default written to the database Postgres which is the standard datastore.

Technical Issues and Architecture

In the following section we present a short overview on the architecture and technical background of the project.

The Geoportal has been realized using the Java Struts-Framework, which is based on Java Server Pages (JSP) and Java Servlets running in a Jakarta-Tomcat Servlet Container. Some parts of the user interface also use Javascript. The database where all the geodata as well the metadata is being stored is PostgreSQL with the postGIS extension in order to allow storing of spatial data in an OGC compliant way as simple features with their respective set of methods for converting, describing and analysing (geometrical/topological) the geodata.

The portal consists of different Web applications:

- The Catalog Service (degree) manages all metadata and allows querying and filtering these in order to find datasets according to the search criteria.
- A transactional Web Feature Service (WFS) (realized through the geoserver project) as well as a Web Map Service (WMS).
- Last there is the portal itself which consists of the client-applications (front-end for the user) that communicates with the different web services (Catalog Service, WFS, WMS) through http-requests (mostly exchanging XML-documents).

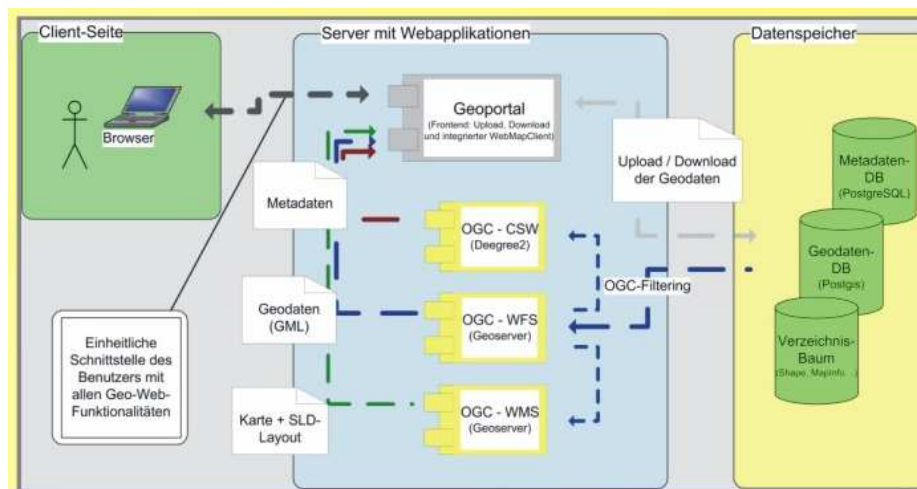


figure 1: UML-diagram of the used components

User Interface and Search capabilities

Dependent from the format of the data that has been uploaded and the converting tools used provided by geoXchange there are different download formats available. In any case the geodata can be downloaded in the format in which they were uploaded and also as ESRI Shape-files and of course as GML as the standard output of the Web Feature Service.

- Search:* allows the search of geodata and webservice.
- Upload:* allows you to upload own free geodata to geoportal-server.
- Download:* allows the user to download geodata
- Metadata:* allows the user to register metadata for the geodata he has uploaded.
- MapViewer:* starts the visualization of free maps and feature-collections.
- Test-Page* allows to test HTTP-requests sending them to a web-service.
- Info:* information about the implementation of geoportal is available
- Contact:* Contact-adresses for feedback and questions.

The user has the possibility to insert the metadata into the database either with graphical user interface or per upload of ISO-XML- or DublinCore-RDF files.

Using the search user interface the user has the possibility to search for all ISO 19115 Core Metadata-elements provided by the current implementation of the portal. Due to limitations in manpower we do not yet support all metadata elements (>300), but only a small subset. We will extend the supported set of metadata elements in future releases of the geoXchange portal. Inserting metadata is always a time-consuming task and in most cases only a minimal set of metadata elements are actually dilled in in most projects. Automated extraction of metadata elements from existing datasets is still a research issue (Manso 2005) and we could not work on this in detail so far.

The following search criteria are available within the current prototype:

Title/keywords

The important criteria is the one with which you can search for words in the title or optional in the abstract, the latter one is a short describing text about the geodata.

Features of the text search:

- Extract terms with a leading minus: -term
- Insert word groups with quotation marks: "word group"
- An automatic search over upper and lower case letters and with the first letter to upper case: if input equals: inPut terms will be searched for: input,INPUT,Input and inPut
- Search in the abstract element optionally
- Search for all words optionally: result contains only metadata sets which have fullfill all the entered word conditions

Category

The user can choose from a predetermined list of topic categories, the corresponding term for the data searched. This list matches the domain elements of the ISO-Core metadata element "Topic category".

Date (of creation, publishing, updating):

The user can select the temporal limitations for his search. Currently this search criteria will only regard the element "Dataset reference date" of the ISO Core metadata set, comparable or rather mapped to the DublinCore element "Date", due to limitations of the Catalog service used. But this limitation will be extended soon to a more exhaustive temporal search.

Location:

At the moment the local search is implemented by receiving the coordinates of the smallest surrounded bounding box (=rectangle)out of a map-display which shows you a world-map to zoom in and out. With every zoom in or with specifying an area of interest (=rectangle as the former described bounding box) new bounding coordinates are displayed on the right side. To activate the local search, set the option field activate to true. Maybe in the future, it could be possible to integrate a Gazetteer Service with which a geographic term like "Berlin" could be inserted in a textfield and then

would be transformed in the same bounding box coordinates as it is done with the world-map. Another feature would be a Thesauri, a synonym dictionary to receive the official geographic name if you enter only a regional known term for a city, river, state and so on.

Conclusion and Outlook

This first version of the portal is going online in 2005 and will be extended until the end of the year to also support raster data (Web Coverage Server) and more extensive search capabilities which include also the search in other OGC Catalogs that are currently being made available worldwide – in Europe most of all inspired through the INSPIRE initiative as an aggregate search within several OGC catalog services, that are being developed in several organizations at the moment. Further the more extensive use of Gazetteer services in combination with Thesaurus Services is planned and we are also watching the work on Web Ontology Services in order to see when there is a good chance to integrate these developments into the portal.

Further we need to raise awareness within the institutions that possibly could contribute datasets on the one hand side and on the other hand side also find existing datasets and include them into our database.

Future extensions include the integration of 2,5 dimensional digital elevation models (DEM) and 3D geodata (3D city models). Work in this direction is already in progress, but with a focus on implementing the basic software components that would be needed for a 3D-geodata infrastructure.

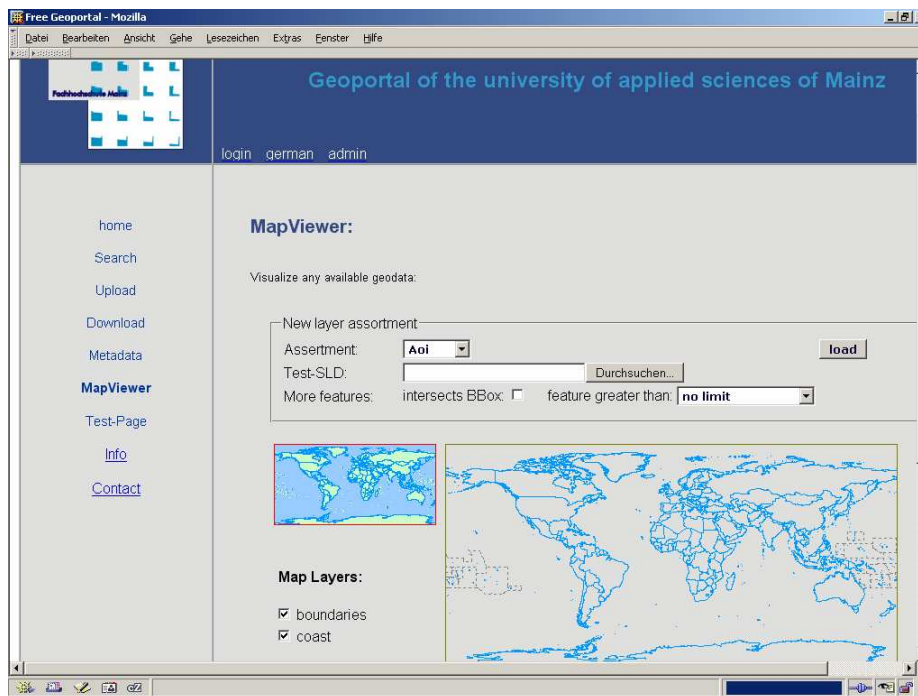
References

1. Baldonado, M., Chang, C.-C.K., Gravano, L., Paepcke, A.: The Stanford Digital Library Metadata Architecture. *Int. J. Digit. Libr.* 1 (1997) 108–121
2. Reuter, A. and Zipf, A. (2005 in print): GIScience - Where Next. In: Fotheringham S. and Wilson, J.P. (eds.): *Handbook to GIS*. Blackwell Publishers. (due Sept. 2005)
3. Zipf, A. (2003): How to enrich GIS development through 'aspects' using formal ontologies. GISRUK 2003. GIS Research UK. London.
4. Zipf, A. and Krüger, S. (2001): TGML - Extending GML by Temporal Constructs - A Proposal for a Spatiotemporal Framework in XML. ACM-GIS 2001. The Ninth ACM International Symp. on Advances in Geographic Information Systems. Atlanta. USA.
5. Noguera-Iso, J. (2005): WEB ONTOLOGY SERVICE, A KEY COMPONENT OF A SPATIAL DATA INFRASTRUCTURE 11th EC-GI & GIS Workshop, ESDI: Setting the Framework Alghero, Sardinia, 29th June - 1st July 2005.
6. Vaccari, L. et al. (2005): A SERVICE ORIENTED APPROACH FOR GEOGRAPHICAL DATA SHARING: 11th EC-GI & GIS Workshop, ESDI: Setting the Framework Alghero, Sardinia, 29th June - 1st July 2005.
7. Manso, M. A. et al. (2005): NEW ADVANCES IN THE AUTOMATIC METADATA RETRIEVAL FROM GEOGRAPHIC INFORMATION. 11th EC-GI

& GIS Workshop, ESDI: Setting the Framework. Alghero, Sardinia, 29th June - 1st July 2005.

8. <http://www.geometa.info/>
9. Vandembroucke, D. (2005): INSPIRE - STATE OF PLAY STUDY: STATUS OF THE NATIONAL SPATIAL DATA INFRASTRUCTURES IN EUROPE. 11th EC-GI & GIS Workshop, ESDI: Setting the Framework. Alghero, Sardinia, 29th June - 1st July 2005.
10. Ostländer, N. (2005): DEVELOPING AN SDI FOR TIME-VARIANT AND MULTI-LINGUAL INFORMATION DISSEMINATION AND DATA DISTRIBUTION. 11th EC-GI & GIS Workshop, ESDI: Setting the Framework. Alghero, Sardinia, 29th June - 1st July 2005.

Attachments: Screenshots of the current geoXchange portal prototype



Free Geoportal - Mozilla

http://www2.geoinform.fh-mainz.de:8095/geoportal/language.do?lang=en

i3 mainz
 Institut für Raumbezogene
 Informations- und Messtechnik
 Fachhochschule Mainz

Geoportal of the university of applied sciences of Mainz

login german admin

home
 Search
 Upload
 Download
 Metadata
 MapViewer
 Test-Page
 Info
 Contact

Geoportal of the university of applied sciences of Mainz

This geoportal supports the search, upload and download of free geodata and metadata.

Search: allows the search of geodata and webservice.
 Upload: allows you to upload own free geodata to geoportal-server.
 Download: here you can download
 Metadata: here you can register the metadata for your uploaded geodata.
 MapViewer: start the visualization of free maps and feature-collections.
 Test-Page: allows you to test HTTP-requests sending them to every web-service.
 Info: here you will find further information about the implementation of geoportal.
 Contact: Contact-adresses for feed-back and questions.

Free Geoportal - Mozilla

Geoportal of the university of applied sciences of Mainz

login german admin

home
 Search
 Upload
 Download
 Metadata
 MapViewer
 Test-Page
 Info
 Contact

Metadata form for geodata description:

On this page you can describe your uploaded geodata with metadata according to the standard ISO 19115.

- With * marked metadata must be entered (elements belongs to ISO CORE mandatory metadata set, rest is ISO conditional/optional)
- With * marked metadata belongs to the Dublin CORE Metadata ElementSet (DCMS)

[Help](#)

MD_Metadata

CI_ResponsibleParty

* Metadata point of contact:

* Date stamp: Day Month Year

File identifier:

Standard name:

Standard version:

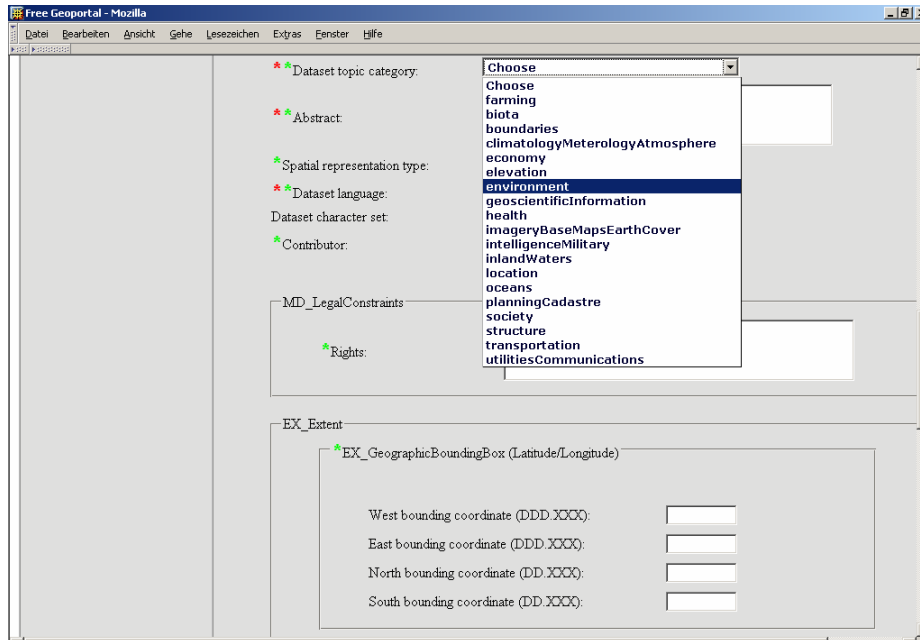


Fig. 1. Example of the User Interface

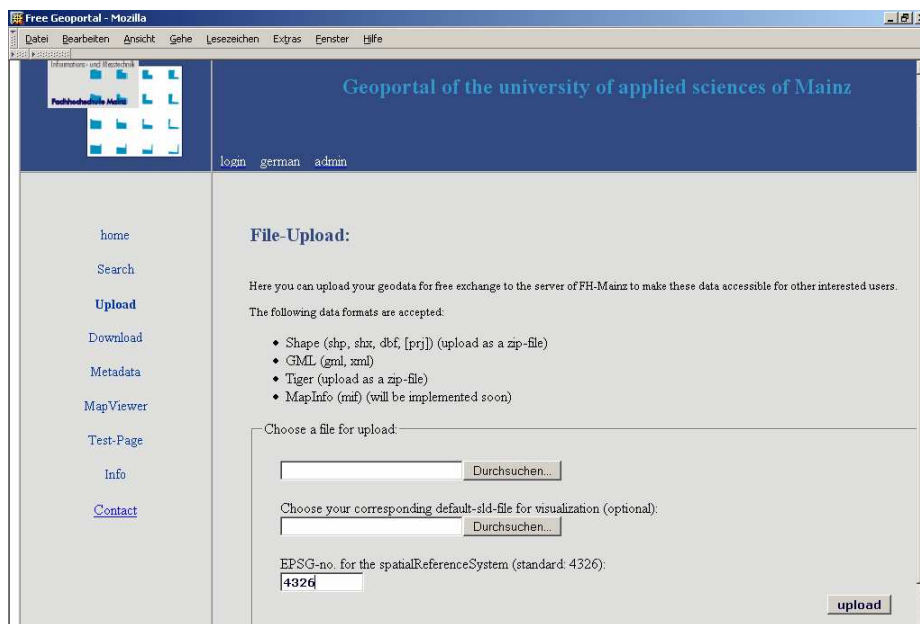


Fig. 1. Example of the User Interface

