CO-328

GEOVITE - A SERVICE-DRIVEN SOLUTION FOR AN ON-DEMAND, USER-FRIENDLY WEB ACCESS TO GEODATA

IOSIFESCU C., IOSIFESCU I., JENNY H., HURNI L.

ETH Zurich, ZURICH, SWITZERLAND

ABSTRACT

The access to geodata is not always straightforward: it often requires time to identify if the geodata we look for is available in the area of interest, to understand the different storage models and the available data formats, to clip the datasets to the desired extent and finally to convert the geodata into the required data format and projection before actually proceeding with the research. This was the case for the ETH Zurich researchers, who were often confronted with such challenges when using geodatasets from the Swiss national mapping agency (swisstopo) in their research.

The GeoVITe (GEOdata Visualisation and Interactive Training Environment) project at the Institute of Cartography, ETH Zurich was started in the pursuit of a viable solution to these problems. Its goal was to provide a clear and intuitive overview of available products with on-demand data processing and download possibilities over the Internet for ETH employees. Our approach consists of a service-driven system for an on-demand, user-friendly Web access to geodata, with a clear separation between data, logic and presentation.

This paper documents in detail the concepts, technologies and workflow used for creating the GeoVITe system. Our experience can be used to recreate the results and to produce similar operational platforms for user-friendly Web access to geodata.

BACKGROUND

The world moves towards on-demand access to information. Large amounts of data are daily provided to users on the Internet, in form of videos on youtube, voice over Skype or satellite imagery over Google Maps. But Web access to information should not only enrich our private life, but it should also facilitate research by providing access to professional vector and raster geospatial datasets with the convenience, simplicity and user-friendliness of the Internet.

Geodata is needed for many research activities in the field of environmental research, architecture, geomatics, cartography and many others. Therefore the ETH Domain receives geodata for internal use at preferential prices from the Swiss national mapping agency (swisstopo). The current official way of distribution to the researchers is a file system located on an internal server. This is organized according to the delivery date and by product name.

Because the data is not free of charge, each user has to specifically ask for the desired files and has to wait till the request is approved. Beside this wait-time, there are other difficulties with which the researcher is confronted.

The first difficulty for the geodata user is to select the desired region. The original raster files come usually as map sheets (e.g. ca. 250 map sheets for a product in the scale 1:25000 for the whole Switzerland). The user has to look on a low-resolution overview map that contains the map sheet numbers and choose the right ones for the desired region. The vector files come in shape-format, one for each defined layer, and therefore it is not possible to select a region; the user will have to handle the whole layer.

In these conditions, after receiving access permission, it happens quite often that the users notice, that the desired region is not completely represented on the requested files and they have to ask for more files.

Another difficulty is when the desired region is at the intersection of more map sheets, which is usually the case. Many end-users do not know how to use GIS programs nor do they need them for their everyday work. But without a GIS program, it is difficult to match the neighbouring map-sheets to each-other. It may also happen that the desired region is smaller than a map-sheet, and it needs to be clipped. On the other hand, sometimes a map sheet may need much storage space, like for products such as swissimage25, the orthophoto-mosaic in a ground resolution of 0.25m, where one file is 600MB large. If the user needs many such files, it may be difficult for a standard graphics editor to handle such large amounts of data and the upper limit of the computing capacity may be quickly reached.

Another challenge is to choose the correct geodata product, to receive an overview of the product types and their spatial availability. Complicated file/ folder names on a file storage system make this difficult for the user.

All these difficulties led to the necessity of developing an easier to use and more straightforward tool for accessing geodata (see also Iosifescu et al., 2010). It was decided that a what-you-see-is-what-you-get system is highly desirable (Jenny et al., 2010).

Our research project started in 2004 and for a long time it was in a developing and testing phase. Eventually it came to be a tool used online by ETH employees and students, with enthusiastic reports from the end-users and a great potential for replacing the file system. In its more than 6-years existence, it profited from many developers, gathered experience in what the users need most, and is now able to offer a stable access to geodata through a user-friendly interface.

APPROACH AND METHODS

First, we had to decide on the system architecture and then specifically what technologies could be used. We decided on a three-tier architecture. The **three-tier architectural design pattern** has three main constitutive components corresponding to the data tier, the logic tier and the presentation tier (Eckerson, 1995). The fundamental characteristic of the three-tier architecture is its linear topology and the clear communication interfaces between components (Figure 1).

In terms of technologies to be used, we need a spatially-enhanced database, where we can store the products and which could be managed by a GIS system. From the beginning, the chosen solution was to offer this tool online. We decided to use services to allow the visualization and download of the geodata. We need a graphical user interface (GUI) that can be customized for our specific requirements. Accordingly, in our case, the data tier is represented by the database, the application tier is represented by the services for visualization and download and the presentation/ interaction tier is represented by the user interface.

We ran the proof of concept solution on one server that contained the database, an ArcGIS Server installation, the visualisation, geocoding and geoprocessing services, and the internet server. That way, we observed that the system was feasible and that the test-users were very content with the gain in time that resulted from the graphical access to geodata.

As the system evolved, we realized that we need a proper architecture also in terms of physical machines, in order to have a stable tool and to offer access to more and more products. This is why we introduced the three-tier-architecture also at the physical organization level. That is, we decoupled the database server from the services servers and from the internet server. Figure 1 shows the current system architecture.

Figure 1. Geovite system architecture: a three-tier-architecture

The three-tier-architecture concept is mirrored not only in the clear separation between the three tiers, but also in the physical organisation of the servers. That is, the database, the services and the Web server are located on different machines connected over the internal network. This architecture allows for more than one database server, which enables a backup and replication of the data-container on separate machines. Additional servers for hosting the services improve the system scalability by allowing more users to be connected simultaneously. An additional Web server acts redundant and minimises the down-times of the platform, thus improving the availability. Therefore, switching to this architecture proved to be an important step further in the development of our system.

Now regarding the specific technologies that we use, we need **a database management system** (DBMS), because of the numerous and large geospatial datasets. For example, one mosaic-dataset can be up to 3.5TB (swissimage25). We could use a commercial relational DBMS such as IBM DB2 or the open-source

PostgreSQL, both are compatible, because we integrate the database into our system using the middleware from ESRI, ArcSDE (Spatial Database Engine), that facilitates the storing and retrieval of spatial data.

To grab this data from the database and offer it to the user we use an **ArcGIS Server** installation. The GIS Server hosts and runs **services**, by means of its server object manager (SOM) and server object containers (SOCs) (ArcGIS Server, 2011). The logic tier located on the GIS Server is represented by the visualization, geoprocessing and geocoding services.

For the most important functionality we offer to the end-user, the download of geodata, we employ Geoprocessing Services. We create these in the following way: we open a desktop ArcGIS Map and we create a new Toolbox. Here, we add a Model. We need for example a Clipping service. We insert as parameters the data from the database and the extent, and then we add the clip-function from an ArcGIS predefined toolbox. We save the result in a temporary folder (%scratchworkspace% in Figure 2) that we create, and we archive this result. The Figure 2. shows the created vector clipping Model.

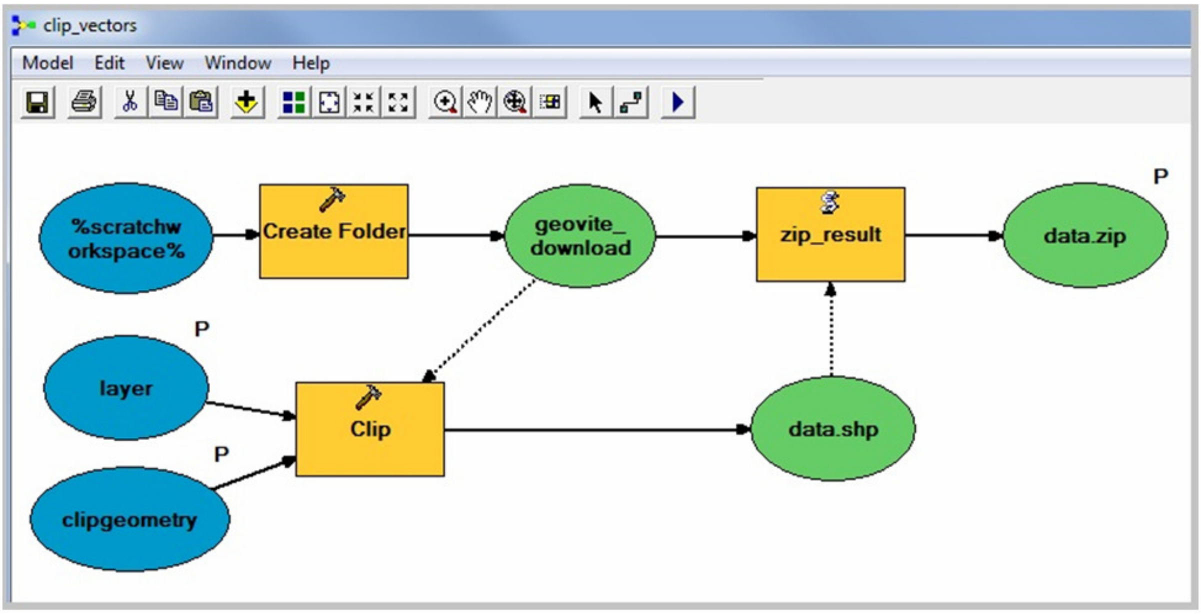


Figure 2. The Model for the Geoprocessing Service clip_vectors; blue codes the input, yellow the imported functions from existing toolboxes, green is the output; P refers to input or output parameters, the arrows with continuous line are the workflow directions and the arrows with interrupted line are conditions (the pointed to will be executed only after the previous finished).

We add to the map the vector layers contained in the database, and this clip_vectors model from our toolbox. We save this as a map (MXD) file. We transfer either the mxd file with the toolbox or just the toolbox itself to the services GIS Server, and from ArcGIS Server Manager we create a new geoprocessing service, based on the previously designed model, which resides in this toolbox. The resulting zip file (e.g. data.zip in Figure 1) is retrieved from the GIS Server and sent to the user, at a request from the GUI.

We created for Geovite also other models such as for clipping a raster (mosaic) dataset, saving a file in a different format (DXF for vectors and ASC for digital terrain models). We also created a geocoding service, which translates location names into coordinates and a reverse geocoding service, meaning the translation from coordinates to a location name in the neighbourhood (if exists).

The WMS visualization services are very important in the graphical user interface; they offer a what-you-see-is-what-you-get feeling. We created a map or an image service from the ArcGIS Server Manager on the GIS Server for every product in the database. As the user chooses in the GUI from the drop-down list one of the map products available, we send a Web Map Service (WMS) request to the server and display the obtained image on the screen. When the user navigates, i.e. changes the product, switches visibility on or off, changes the location, zooms in or out, we send WMS requests to the server and display each time the updated map.

WMS is a standard protocol for serving georeferenced map images over the Web that are generated by a GIS server using data from a spatially enabled database. Maps are provided with the GetMap request and metadata with GetCapabilities and GetFeatureInfo requests (ISO, 2005).

Our **graphical user interface** (GUI) is encapsulated in Java Servlets and JSP technology and hosted on a Tomcat Servlet Container (reference Tomcat) on a Web Server, and therefore available over the Internet.

For the GUI itself, our presentation/ interaction tier, we use JavaScript and an SVG framework (Neumann et al., 2003), that we further developed and adapted to the needs of our project. The user interface interacts with the visualization and download services by using a simple request-response mechanism. From the GUI we let the users choose the products/ layers they need and the clip-geometry: the extent, or what we call the dataframe. We draw this dataframe on the screen using SVG (scalable vector graphics).

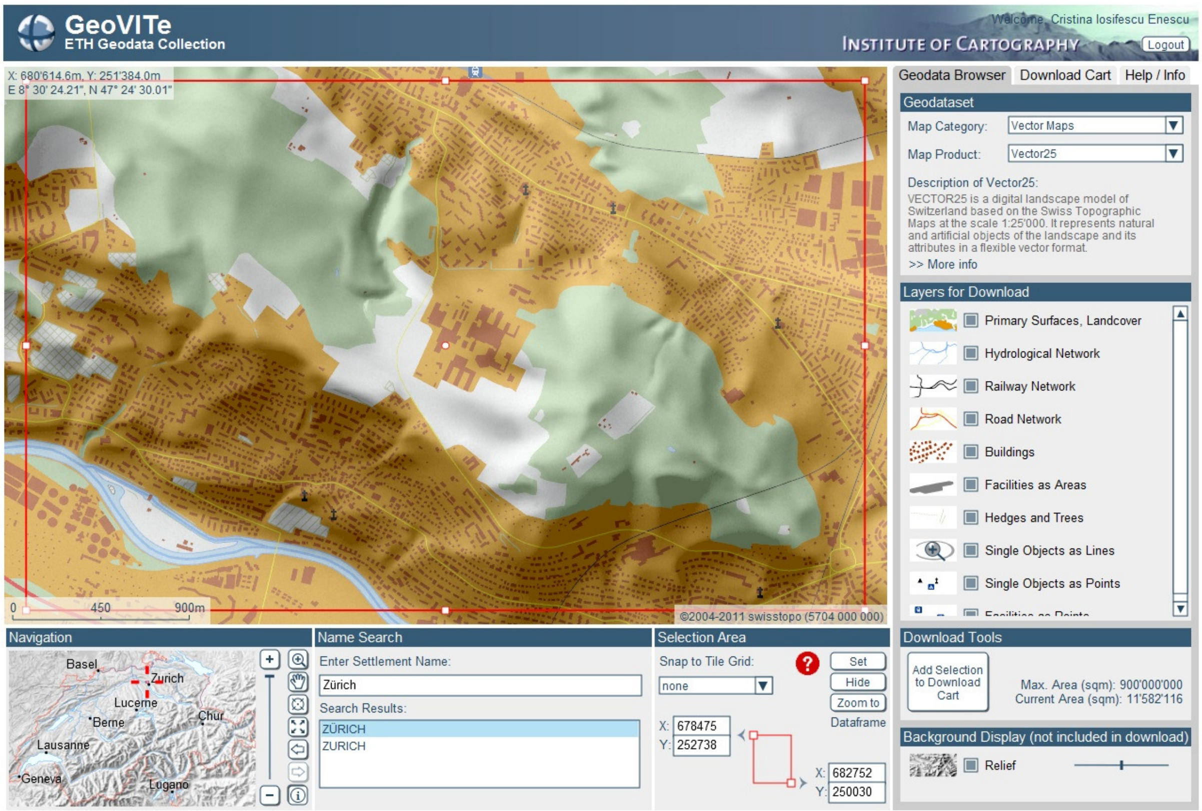


Figure 3. The Geovite Graphical User Interface

The spatial navigation controls are the elementary interactive components allowing zooming and panning the map. These controls are automatically supported by the SVG framework. The map navigation enables the user to choose an area of interest, to zoom in or out, pan, recenter, obtain full view or to go back and forth to previous views of the map. A zoom slider allows for continuous zooming. The reference map allows to reposition the map extent and to get local-global orientation cues.

The layer manager shows and hides map layers. It allows the user to select and deselect them with a checkbox from a list. It also contains a short description about the map category of the displayed layers and provides a legend for each layer (Figure 3).

The users can add the chosen products to the download cart. Here, they may change the format of the data to be downloaded (e.g. GeoTIFF, ASCII Grid, ESRI SHP, Autodesk Drawing Interchange Format DXF) and the projection (e.g. CH1903/LV03, WGS84, UTM 32N (ED50), UTM 32N (WGS84)); this will add parameters to the triggered geoprocessing service, when the user press on the generate file button.

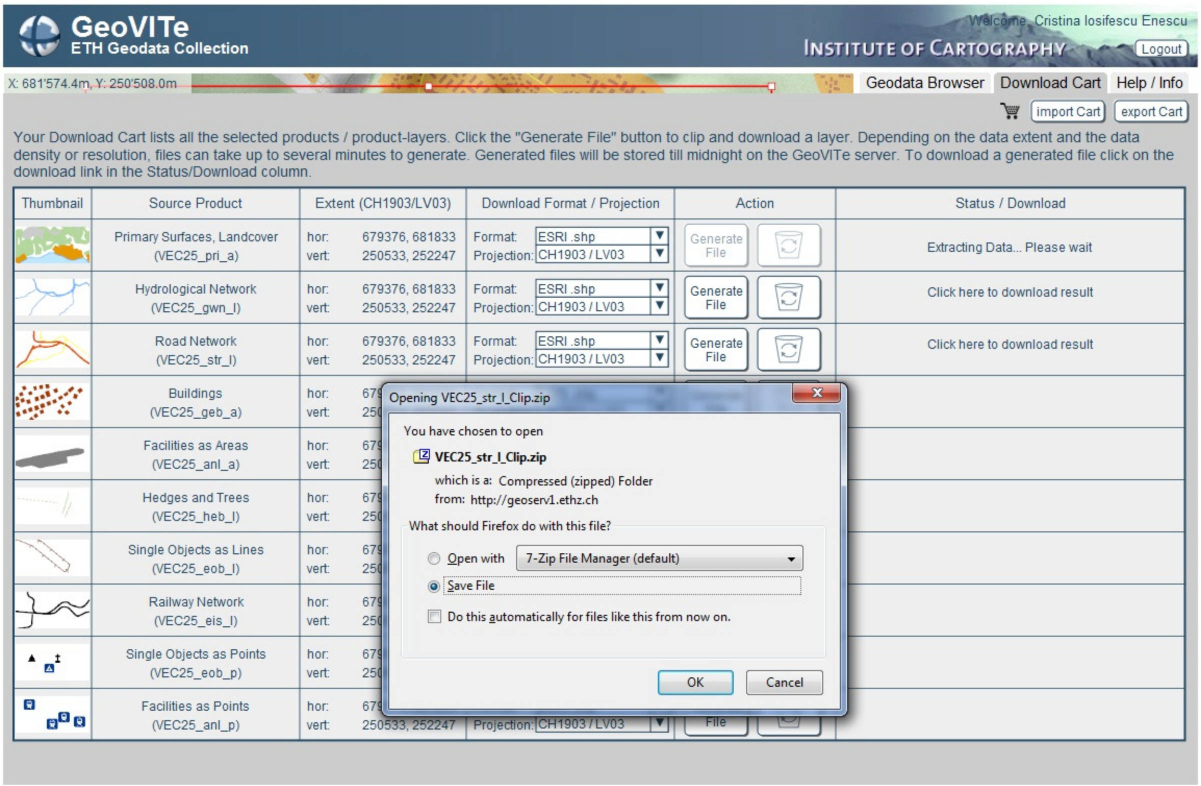


Figure 4. Geovite – the Download Tab

After a waiting time between a few seconds and a couple of minutes (depending on the time necessary to the geoprocessing service to execute the requested operation) the link to the result is displayed in the download table (Figure 4). The user will save the file on the desktop computer and can immediately use it. Consequently, the time in which the user comes to the desired product is drastically reduced, compared to the file system.

Now, that we presented the basic component of our system, we want to picture some other aspects, that are important for our project.

In order to offer the geodata through Geovite, we need to follow some preparational steps. After we receive the original data (updates come every quarter year from swisstopo), we organize the products by the production year. Then, for the raster products, we proceed with a mosaic of the map-sheets of a year and we update the “latest” version of the product. We upload the mosaics, respectively the vector data on the database server through ArcSDE. After that, we create the visualization services for each product, defining the scale to which the products are visible, and we represent the vector products cartographically. The clipping services are created as we mentioned above for one raster/ vector dataset and then other datasets are added to these.

The details of the products offered by Geovite and their structure may be easily changed from an XML File, where also the visualization and geoprocessing services requests are referenced in clear text for each dataset.

We paid attention also to specific requests. Because only ETH staff may download geodata, but students also need geodata for their projects and they know best which region, or what product they need, we provide a special service for the students and their supervisors: the download cart. The students may log in into Geovite, select the area and the product they need, and then save the cart and send it to their supervisor. In this way, we can verify who downloads what products, and the supervisor is responsible for the data usage.

Many users of our system are from the department of architecture. They have specific needs, for example the contour lines at a certain interval in a vector format. This is not provided as such in the original data from swisstopo. Considering this need and the fact that only a GIS program (that architects do not use in their everyday work) is otherwise able to generate this data, we went beyond the original data, and we provide also the possibility to generate contour lines on request from the digital terrain model, transparent to the end-user, through a geoprocessing service, created similar to the clip_vectors previously described.

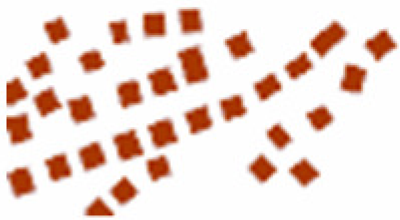
In addition to the visualization, which is more common on the web, we introduced the on-demand user-friendly download possibility. This is our great innovation and it was the main demand to Geovite, the purpose for which it was developed.

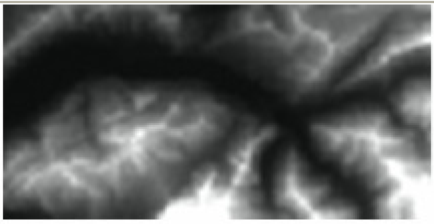
RESULTS

The main functions of Geovite are the visualization and navigation through geodata and the download of datasets. We have at this time an operational system in a beta version that supports already a great amount of geodata and a fair number of users simultaneously.

As our system became more stable, more products were added to our database and it became clear that we also needed a geodata history, organized by the year of publication for the topographical maps, respectively by the year of flight for the orthophoto products.

We would like to present in the Table 1 some of the most important datasets that are currently offered through Geovite to ETH employees.

Map Category	**Example description of a product**	**Example thumbnails**
Vector Maps: Topographic maps in the scale 1:200000 and 1:25000 with layers such as primary surfaces, administrative boundaries, hydrology, buildings, road network, railway network; Toponymy at scales from 1:500000 to 1:25000	VECTOR200 is a digital landscape model of Switzerland based on the Swiss Topographic Maps at the scale 1:200000. It represents natural and artificial objects of the landscape in a flexible vector format with a high generalisation grade.	 
Raster Maps: Topographic maps in the scale from 1:1000000 to 1:25000 with or without relief, with the production year from 2010 back to 1938, produced directly digital (PK) or scanned (LK); Cartographical raster masks for the previous listed scales with raster layers such as borderline, lakes, cities, situation, text, highway network, etc; Relief	PK200 is a raster version of the Swiss Topographic Map 1:200000 combining several layers in one map. This product is useful as generalized background map or for planning, communication or market research.	  

Digital Elevation Models: Digital Elevation Model with a 25m resolution, Digital Terrain Model and Digital Surface Model with a 2m resolution	The digital terrain model of the cadastral surveying (DTM-AV) represents the surface of the ground without vegetation and buildings. The altimetric precision (simple standard deviation) of the terrain model on an unspecified site is better than 0.5 m. The density of point is approximately 1 point by 2 m2 in open areas.	 
Airborne Orthoimages: Landsat25 with a 25m resolution, Spot with a 5m resolution, Swissimage with a 0.5m, respectively 0.25m resolution	The orthophoto mosaic SWISSIMAGE is a composition of digital color aerial photographs. Ground resolution: 0.25m. Swissimage 25 is available for the Non-alpine regions of Switzerland.	 
Thematic Maps: Geological Atlas at the scales 1:500000 and 1:25000 in vector and raster format, Tectonic Map and Hydrogeological Map at the scale 1:50000, vector and raster, The Last Glacial Maximum, The Gravimetric Map at 1:500000	The Tectonic Map of Switzerland 1:500000 (raster) displays the large-scale structural and tectonic units, which contribute to a better understanding of the geological buildup of the Alps.	 

Historic Maps: Dufour Map, first and last edition at the scale 1:100000, Siegfried Map first and last edition at the scales 1:50000 and 1:25000	Siegfried Map, First Edition 25 is a historic raster map in the scale 1:25000. The map tiles were published between 1870 and 1922 and show only the plain parts of Switzerland.	

Table 1. Datasets available through Geovite

Obviously, several challenges had to be overcome, considering the huge and heterogeneous amount of data that had to be managed. The geodatasets are very different in terms of their date of production, their type, scale and spatial resolution, as you can see in Table 1.

CONCLUSION AND FUTURE PLANS

Cartographers, geomatic engineers and computer scientists at our institute were employed in finding a viable solution for user-friendly geodata access and distribution.

In addition to the proof of concept phase, where we tested the technology and tuned it for our needs, we were challenged with the huge amounts of data and with a system that had to allow a bigger number of users simultaneously choosing and downloading geodata. Some of the problems were solved, but the need of having a back-up and offering a more reliable and scalable service based on an extended infrastructure is to be considered by the fully operational Geovite.

We are confident, that the research done in the frame of our project has already led to a feasible solution and we look forward to further contribute to improving geodata distribution at ETH Zurich.

REFERENCES

ArcGIS Server (2011): Website of ArcGIS Server, Environmental Systems Research Institute, www.esri.com/software/arcgis/arcgisserver/, accessed on 2011-02-15.

Eckerson, Wayne W. (1995). Three tier client/server architecture: Achieving scalability, performance, and efficiency in client server applications. In: Open Information Systems 10 (1), available through http://www.mendeley.com/.

Iosifescu I. and Hurni L. (2010). GIS Platform for Interdisciplinary Environmental Research, in Proceedings of the 7th ICA Mountain Cartography Workshop, Borsa, Romania (in press)

Jenny, H., Neumann, A., Jenny, B. and Hurni, L. (2010). A WYSIWYG interface for user-friendly access to geospatial data collections. In: M. Jobst (ed.), Preservation in Digital Cartography, Lecture Notes in Geoinformation and Cartography, Springer, p. 221-238.

Neumann, A. and Winter, A.M. (2003). Vector-based Web Cartography: Enabler SVG, http://www.carto.net/papers/svg/index_e.shtml, accessed on 2011-02-15