How do Free and Open Geodata and Open Standards fit together?

From Sceptisim versus high Potential to real Applications.

Michael Auer, Alexander Zipf
University of Bonn, Department of Geography, Chair of Cartography
http://www.geographie.uni-bonn.de/karto/
{auer, zipf}@geographie.uni-bonn.de

Abstract
Currently we encounter two trends within the broader GI community: one is the move to web-based applications based on open standards as defined by the Open Geospatial Consortium (OGC) resulting in more flexible and interoperable solutions. The other is the raise of the prosumer-oriented GeoWeb2.0, with its user generated content. Goodchild (2007) gives an overview of these global collaborations with respect to geographic information. He calls the phenomenon Voluntary Geographic Information (VGI). One of the most striking examples of VGI is the OpenStreetMap (OSM) project. It aims at creating and collecting free vector geodata covering the whole planet. While the quality is still heterogeneous (Haklay et al 2008), we see already potential for real applications using this data. This paper will point out, why VGI and open standards fit together by complementing one another. We will have a look on the problems of using heterogeneous data in terms of pre-processing and show some possibilities, how to use the potential of VGI and other free geodata in combination with OGC-Standards.

Free and open geodata in combination with open standards

In most European countries geodata is provided either by public or commercial institutions. On the one hand this geodata claims to be of high quality in the sense of accuracy, actuality and homogenous integrity. On the other hand this kind of data tends to be very expensive and therefore isn't suitable for many private or scientific initiatives with the intention to provide new kinds of innovative services in the broader field of geoinformation and communication technologies.

Although there is a broad community of engaged people interested in providing open source software for all kinds of geodata processing, analyzing, visualization and distribution, the use of those creative masses for the development of innovative applications often is limited by the lack of affordable data. Here comes the recently often discussed volunteered geographic information into play. Is VGI an alternative to fill software with data and thus turn it into useful applications? The answer to this question cannot be unambiguous. There may be critical requirements concerning the data quality, e.g. in security-domains, but in many cases, yes, it can be an option as shown later in this paper by some examples. Each application has its own needs of data quality, so you have to weigh the pros

and cons of using VGI particularly with regards to the efforts which have to be made in pre-processing the data to achieve the desired data structure, homogeneity etc which would be necessary for the provided services.
Now what has the use of VGI to do with use open standards? Having in mind the processing hours which the data has to go through in many cases for special preparation, due to heterogeneity or non-standard data structure, as shown later, it becomes pretty clear that the reuse-ability of this data is crucial to achieve efficiency and generate added value – in particular for the professional world of GI-businesses and institutions such as local government. And this is exactly the domain where open standards play an increasing role. The interoperability obtained by using open standards is the key to easily provide multiple applications with different combinations of once prepared VGI-data.

Problems and possibilities with free geodata and open standards
As hinted above there are certain problems in using VGI, but with the use of open standards and the appliance of intelligent pre-processing-algorithms these problems can be minimized to maximize possibilities.
Currently we investigate the quality of OSM in a systematic approach for some specific applications (e.g. landmark extraction for pedestrian routing, suitability for different types of routing – from pedestrians to bicycles and even wheelchairs), try to improve the data quality through algorithms (e.g. the case of addresses in geocoding), homogenize the data for visualization and other purposes in 2D (OGC-WMS/WFS, see http://www.osm-wms.de), spatial POI search for different categories (OpenLS Directory Service)and even3D (OGC-

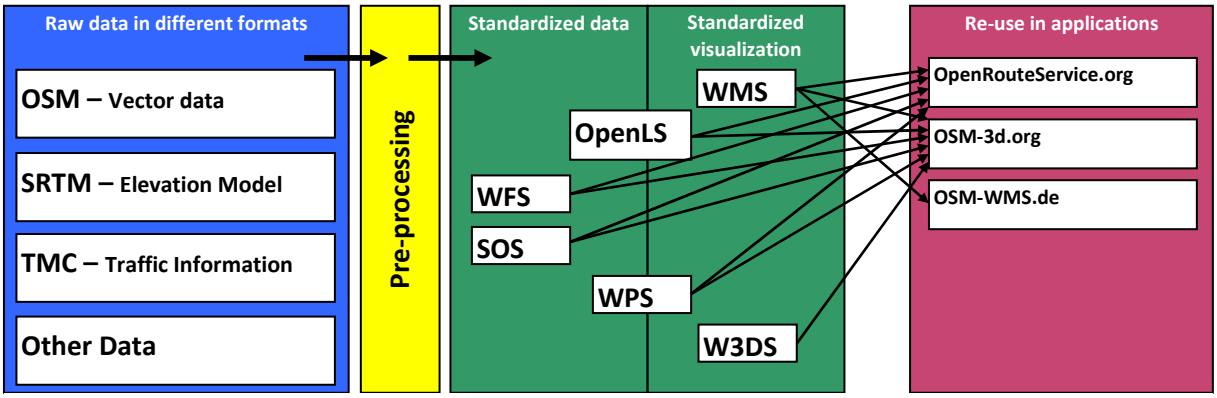

Fig. 1: Multiple re-use of once processed and standardized data

W3DS) and combine and convert OSM-data with other free data like SRTM-DEM-data to realize even more innovative applications (e.g. www.osm-3d.org).

A first example of an internet application using this data beyond mapping is OpenRouteService.org. OpenRouteService uses data from OpenStreetMaps and open standards of the OGC and delivers a set of location-aware services for the whole of Europe. For LBS the OGC has specified a set of services within the Open Location Services initiative (OpenLS). So far we have implemented five of these:

- Location Utility Service (Geocoder and Reverse Geocoder) ,

- Presentation Service,
- Directory Service (POI-Search) and of course a
- Route Service and recently the new
- Tracking Service.

Further a WMS, WFS and several processes of a OpenGIS Web Processing Service (WPS) are used at OpenRouteService.org in order to deliver specific functions (such as integration with TMC data or other dynamic sensor data via a OGC Sensor Observation Service (SOS), cmp. Mayer & Zipf 2009, Mayer et al. 2009).

Through the integration of OSM data with free digital elevation models even more innovative applications can be realized:

Schilling et al. (2008) show how to integrate free OSM data with the open source Shuttle Radar Topography Mission (SRTM) data to construct a digital elevation model for Germany. That is integrated into the 3D (geo-)spatial data infrastructure (GDI-3D) and used for 3D visualizations on the Web through the OGC Web 3D Service (W3DS) draft specification, for which a new version of the draft specification is currently being developed within the DSS task of the OGC OWS6 testbed. The result is a 3D application for a large area (first case: whole of Germany) based completely on collaboratively collected free geodata and open standards using several OGC services. The service is available online as a first prototype at www.osm-3d.org.

Fig. 2: Screenshot of OSM-3D.org showing Munich based on open data (OSM & SRTM) and open standards (OGC Web Services)

To give an overview of the potential of re-using data through open standards the following table shows processing times for different data preparation steps for our services (numbers as of May 2009):

Process	**Processing Time**	**Service/open Interface**	**Used in Application**
Data conversion: OSM (Europe) to Postgis	27 h per week (10 Mio. Features)	WMS, WFS, OpenLS - DirectoryService	www.osm-wms.de www.osm-3d.org www.openrouteservice.org www.heidelberg-3d.de (Directory Service)
Data Conversion: OSM (Europe) to RouteGraph2D	1 h per day	OpenLS - RouteService	www.openrouteservice.org www.osm-3d.org
Data conversion: OSM-buildings (Germany) to VRML to Postgis	4.5 h per day (ca. 330000 Features)	W3DS	www.osm-3d.org
Data conversion: OSM-POI's & Places(Germany) to VRML to Postgis	8h per day (ca. 660000 Features)	W3DS	www.osm-3d.org
Data conversion: OSM-Streetlabels (Germany) to VRML to Postgis	1 day (ca. 1.2 Mio.)	W3DS	www.osm-3d.org
OSM (Europe) to AdressDB	4 h per week	OpenLS – Location Utility Service (geocoder/reverse geocoder)	www.openrouteservice.org www.osm-3d.org
RouteGraph2D (Germany) + SRTM to RouteGraph3D	2 days	OpenLS – RouteService(3D), WPS(HeightProfile)	www.osm-3d.org www.openrouteservice.org
OSM (Germany) + SRTM to TIN	1300 h/CPU	W3DS	www.osm-3d.org

As you can see in the table above, every reuse of processed data can save a lot of processing-costs (time) while working with heterogeneous and frequently changing data like OSM-data.

All these processes need permanent development to cope with changing input data and to increase the output data quality by certain algorithms. Therefore further research is in progress, such as OSM-Quality-Evaluation or User-Interfaces for Web2.0-driven Quality control and enhancement.
VGI contains a lot of local specialist knowledge about our physical and non-physical environment, which forms a great potential for innovative and useful services. The challenge is to find out ways to systematically extract that information and make it usable to many others by provision through open standards. A big benefit is that these standards offer very modularized functionality that can easily be reused in different applications. This has been shown in several projects mentioned above and the even more dynamic chaining of standardized services through technologies like BPEL (Business Process Execution Language) or even OGC Web Processing Services (WPS)(Weiser 2007, Stollberg & Zipf 2009). A new research activity that will use open standards, open source and free geodata in combination based on and extending the above mentioned services will be an application of evacuation simulation similar to earlier work by Haase et al. 2008.

Conclusion

In the beginning of the project OpenStreetMap most professional and scientific users of geoinformation expressed mostly skepticism about the value of user-generated geocontent (and some still do). Recently this changed to some hype about this kind of data - and community-based approaches to collect geodata in general. For example also professional data providers from Google to TomTom try to exploit prosumer-based approaches by providing means to have users contribute information to their existing professional geodata sources. These artificial "communities" seem not to work as well as the voluntary ones though. A lot of small enterprises now try to use user-generated geodata for different applications – mostly providing maps for specific purposes. In contrast to that we could realize a range of real application and value-added (web) services in different domains and provide those for large regions differing in scale from whole nations (Germany) to Europe. Our experiences show that free and open geodata and open standards fit well together in several points. Open geodata, like commercial data is expensive, either in processing costs or in monetary costs. In both cases open standards can help to reduce those costs by enabling an easy recombination and re-use of data and services into multiple applications. Of course there are issues related with the ease of use of OGC standards – they tend to be more complicated than the APIs of many Web 2.0 projects, which tend to use technologies such as REST (Schmitz et al. 2009). But the big advantage of the existing (OGC) standards is, that they are well established in the professional GI community and therefore supported by a wide range of tools and applications and are taught in universities (also because of their independency from companies).
Despite disadvantages in developing time free datasets, like OSM, are helpful in research, as applications can be tested with really large datasets to check

performance and stability. Another point, the challenge to cope with ever changing and inhomogeneous datasets leads to new methods of data-quality-evaluation and enhancement methods.
A further point, which should not be underestimated, is the community, which tends to arise around a VGI-Project like OSM. Such a group of prosumers has an intrinsic interest in what is done with 'their' data. In this context, on the one hand, they can act as a vivid group of software and usability testers by giving feedback. On the other hand they build a large user-group where news of new developments and new ideas can be spread out sometimes very fast and easily.

References

Goodchild, M.F. (2007): Citizens as sensors: the world of volunteered geography, In: GeoJournal, Vol. 69. Issue 4, pp. 211-221.

Haase M., A. Zipf , P. Neis , V. de Camargo (2008): Interoperable Routing Services in the Context of Evacuation Schemes due to Urban Flooding. EnviroInfo 2008 Conference. Environmental Informatics and Industrial Ecology. Lueneburg, Germany.

Haklay, M. (2008): How good is OpenStreetMap information? A comparative study of OpenStreetMap and Ordnance Survey datasets for London and the rest of England, in: Environment & Planning (under review).

Mayer, C. and Zipf, A. (2009): Integration and Visualization of dynamic Sensor Data into 3D Spatial Data Infrastructures in a standardized way. GeoViz 2009. Contribution of Geovisualization to the concept of the Digital City. Workshop. Hamburg. Germany.

Mayer, C., Stollberg, B. and Zipf, A. (2009): Providing near Real-time Traffic Information within Spatial Data Infrastructures. The International Conference on Advanced Geographic Information Systems & Web Services (GEOWS 2009). Cancun. Mexico

Neis, P., Zipf, A (2008): OpenRouteService.org is three times "Open": Combining OpenSource, OpenLS and OpenStreetMaps. GIS Research UK (GISRUK 08). Manchester

Neubauer, N., M. Over, A. Schilling, A. Zipf (2009): Virtual Cities 2.0: Generating web-based 3D city models and landscapes based on free and user generated data (OpenStreetMap). GeoViz 2009. Contribution of Geovisualization to the concept of the Digital City. Workshop. Hamburg. Germany.

Schilling, A.; Lanig, S.; Neis, P.; Zipf, A. (2008): Integrating Terrain Surface and Street Network for 3D Routing. 3D Geoinfo 08. 3rd International Workshop on 3D Geo-Information. Seoul. South Korea.

Schmitz, S., Neis, P. and A. Zipf (2008): New Applications based on Collaborative Geodata - the Case of Routing. Submitted for: XXVIII INCA International Congress on Collaborative Mapping and SpaceTechnology, Gandhinagar, Gujarat, India.

Schmitz, S., Zipf, A. and Pascal Neis (2008): Proposal to define common resources for OpenGIS Location Services. 5th International Symposium on LBS & TeleCartography. Salzburg. Austria.

Weiser, A., Zipf, A. (2007): Web Service Orchestration (WSO) of OGC Web Services (OWS) for Disaster Management. GI4DM 2007. 3rd International Symposium on Geoinformation for Disaster Management. Toronto, Canada.