PROJECT OPENSTREETMAP AS OPEN AND FREE SOURCE OF GEODATA AND MAPS

Daniel, BARTA[1]

[1]DHV CR, spol. s r.o., Černopolní 39, 618 00 Brno, Czech Republic
daniel.barta@dhv.com

Abstract

The article characterizes the fundamentals of the project OpenStreetMap (OSM), which is an original source of open and free geodata. It describes the current state of the project, its usability for geoinformatics technology (GIT) and anticipates its near future. It is compared to similar free and open source projects. As an example, there is given a focus on the representation of the Czech Republic and the local OSM community.

In the first part, there is a description of the project core and associated external community web services. Very interesting is its starting point, the users' motivation and their mapping capabilities like quality, quantity and actuality. OSM is using a specific XML data format and an easy 2D model of geodata. The data are published under a free license. Only a few rules for editing are required, but many recommendations are given. Specific renderers are used for generating thematic maps.

Geodata and maps are importable to traditional GIT and a extension for editing the specific OSM format is developed.

The next part is focused on the Czech Republic. There is a list of externally imported data, basic data sources and completeness layers. Demos of several thematic maps like highway/street map, cycle map, public transport map and other are shown.

It discusses problems concerning the geometric part of geodata - its topology, descriptive part of geodata and semantics. Further are discussed categorizing objects, area completeness and thematic catch of mapping.

The last part is about the project's future. The near future of the project consists in changing the license. The long-term future will be based more on users' than developers' ideas.

Keywords: OpenStreetMap, geodata, crowdsourcing, open and free geodata

INTRODUCITON AND CONTEXT

OpenStreetMap (OSM) is the name of a community project, which's goal is creating and cumulating free and open source geodata. At a glance, it is possible to use this easy analogy:

- public editing + encyclopedia + free = Wikipedia
- public editing + map + free = OpenStreetMap

From the Table 1, which shows founding years of free and open source projects (FOSS) in the field of GIT/IT, it is possible to see, that firstly the idea of FOSS and an appropriate license was formulated. Afterwards, the development of traditional specific software, focused on a small group of specialists, starts. Next comes the development of applications for home and office usage and subsequently it aggregates the creative potential of its consumers (i.e. SETI@home, Wikipedia, OSM).

This evolution comes as an effect of progress in technologies, as military and space survey high technologies (i.e. computer, GPS, internet), which became a part of the daily life of people during the past decades. At the end of this evolution, after 2004, comes a specific phenomenon of the Web 2.0 with the involvement of its users in the process of creation of the project. One of its characteristic features is crowdsourcing, which means in the newspeak, the redistribution of a share of the work to the community [1].

Presumptions for the functionality of OSM, respectively collaborative mapping, are [2]:

- Absence of user-friendly maps (price, license or technology)
- Low cost and accessible hardware (i.e. PC, GPS receiver), software and services (Internet, GPS)
- Human needs of creativity and its sharing on the free principle (do it yourself)
- Well fulfilled human rights, freedom and free time

Traditionally, mapping is the most costly part of GIT [3]. The rise of OSM is a reaction to the state, that many base maps all over the world are unavailable for ordinary people due to licenses or financial reasons, although many maps are created from public finances [4]. In the Czech Republic, there is an additional problem: many datasets are available for free access or downloading, but they lack appropriate licenses, which would explicitly define the level of freedom for the usage (i.e. at the the organizations ČUZK, ČSÚ, CENIA) [5]. The situation in Slovakia is similar [36].

Positive examples are maps and data created for the government of USA, licensed exclusively under Public domain (i.e. NASA, US Geological Survey, US Census Bureau).

Following users' motivations are technological limitations of actual maps (i.e. paper map, e-paper map, area limited, actuality, accuracy) or the absence of public maps or geodata.

Table 1. Selected free and opensource IT, GIT and community projects.

Starting year[6]	Project	Focused users	Community goal	Note
1983	GNU Project	public	no	long way to end users
1989	GNU/GPL (General Public License), MIT licence, BSD license	IT developers	no	
1990*	GIS Grass	skilled GIT users	no	
1991	GNU Linux	public	no	long way to end users
1995*	PostgreSQL	IT users	no	
1996*	GIMP	public	no	
1999*	Inkscape (first Sodipodi)	public	no	
1998*	Mozilla (later Firefox)	public	no	
1999	SETI@home	public	crowdsourcing	no creativity
2000*	OpenOffice.Org (first StarOffice)	public	no	
2001	Postgis	GIT users	no	
2001	Wikipedia	public	collaboration	limited to articles
2002	Quantum GIS, JUMP GIS	beginner GIT users	no	
2004	OpenStreetMap	public	collaboration	
2009	ODC/ODbL (Open Data Commons/Open Database License)	public	no	

* project is a fork of FOSS

GEOINFORMATICS REFERENCE AND CORE OF PROJECT

OSM uses only a 2-dimensional geographic frame reference WGS84 (EPSG:4326). In its vector model space it is possible to use the following role and elements [7,8]:

- Node – point (or vertex) with spatial coordinates
- Way – referenced on the nodes
- Area – a way with a specific attribute (at least one from many designated for the area)
- Relation – aggregates elements and describes the relations of one to another

All OSM data, including the history of editing, are saved in PostgreSQL database (db size 1,4 TB) and exchanged and exported in a self-made geodata format *OSM XML format*. The actual version of the last state of the database without history (product named planet) is every week exported into OSM XML format and compressed by bzip2. Its size is 12 GB, every-day changes in OSM XML are compressed by gzip and have cca 50 MB XML. Aside of the model space are appended users' GPS tracklogs in format GPX [9,10,11].

Fig. 1. Czech Republic on the webmap (Mapnik renderer) [31]

The project uses the license *Creative Commons Attribution-ShareAlike 2.0* (CC-BY-SA), but now, it is in migration to a new license *Open Database License* (ODbL). The old license was targeted for multimedia arts; the new license will assure more openness (i.e. creating map compositions from free and license-limited data) with better legal safety of geodatabase. For service software the community it uses mostly the license GNU/GPL [12].

OpenStreetMap Foundation (OSMF) was founded in 2009 in order to support the project. Money will be used for financing core hardware and promotion. OSMF will support also the OSM meeting *State of the map*, which happens every year since 2007 [13,14].

The think-tank of project resides in the UK, where OSM was founded in 2004 by Steve Coast, and in Germany, where it became popular in a short time. In both countries reside corporations, which use OSM as a platform for offering commercial services and products: Cloudmade, ITO World, Geofabrik GmbH.

MOTIVATION AND USERS

Users are the engine and the consumer of project, often at the same time. Specifics of the collaborative mapping are:

- Local knowledge (actual and verifiable – see Figure 6.)
- Number of users (theorem: what 1 man does in 100 days, 100 men do in 1 day)
- Liberty in the theme of mapping (minimal rules, often only recommendations)

For community users is typical a low level of organization and uniformity [15], but a high spontaneity, euphoria and WOW effect. Users have different visions, goals of mapping and methodology. This attitude causes many misunderstandings, mistakes and imperfection. In this collaborative creating, there is expected a positive motivation for any editing and motivation for fixing obtrusive errors [16,17].

Problematic is to estimate the quality assurance, like spatial and attribute completeness, accuracy, depth of detail and the mapping scale. Spatial completeness in OSM mapping is characterized by the density of inhabitants, which means that the best mapped are urban centers and estates [18].

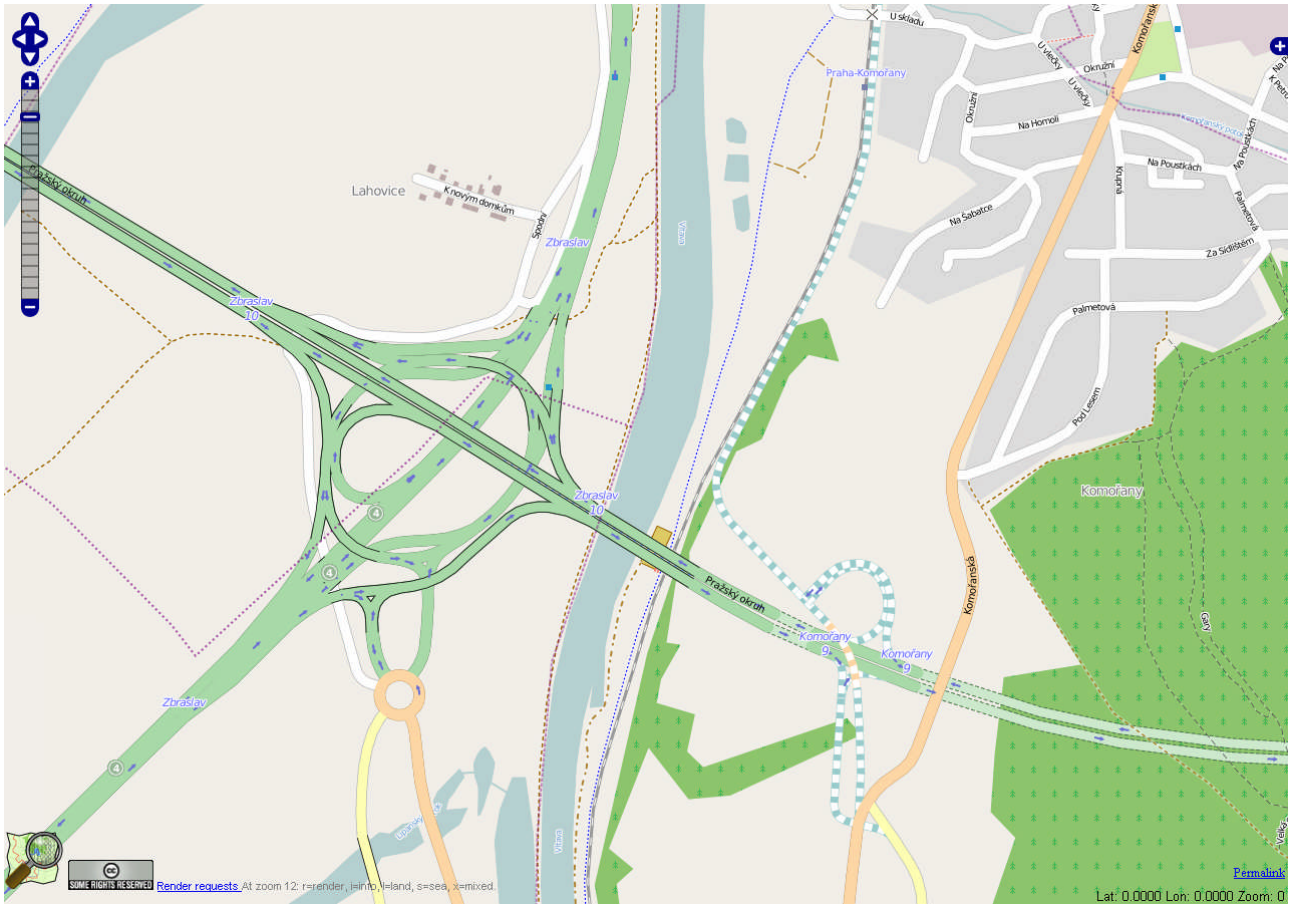

Fig. 6. Flyover junction near Zbraslav and a proposed link to Lahovice on the newly constructed expressway R1, the Prague by-pass. The Prague by-pass has been a part of OSM since 2009. Just on the official opening day, 20. September 2010, its tag has changed from „highway in construction“ to „routable way“. (Mapnik renderer) [38]

In a quantitative evaluation of users we have 300.000 registered from the whole world. Only 5 % were active in the editing of the map in last month. Only 5 % are authors of more than 10 000 nodes (See Table 2).

From the user inquiry it went out that almost all users are men, irrespectively of age groups. 3/4 of users have a university degree and 3/4 of users are working. From the employed almost 3/4 work in commercial sphere. Almost all do the OSM mapping at home, 1/4 of users do it at work or from a mobile device [22].

Main motivations for users are the *project goal*, *altruism* and *activities in local neighborhood*. About 1/4 of users have experience with GIS, but a majority of them is in the low-activity group.

Table 2. Statistics of OSM users and data [19, 20, 21, 22].

Criterion	World	Czech Republic	Great Britain	Germany
Number of nodes	810 million	17 million	19 million	49 million
Number of authors of nodes in the actual version	105 068	2 128	9 482	33 767
Number of registered users	310 000	2 930	-	-
Number of authors of more than 10 000 nods	5%	3%	-	-
Number of authors of more than 100 000 nods	0,6%	0,7%	-	-
Number of active users in Sept. 2010:	5%	-	-	-

DATA AND MAPS USEFUL FOR GIS

The internal data type *XML OSM format* of the project is unique and unknown for traditional GIS software. The OSM geodata model includes these primitives: a point and a polyline. A polyline may become a polygon by adding an attribute to it. Since 2007 API version 0.5 was released, extended by a new primitive: *relation*.

Simplified XML file of OSM data with all three primitives:

```
<?xml?>
<osm version="0.6">
  <node id="" lat="" lon="">
    <tag k="" v="" />
  </node>
  <way id="">
    <nd ref="" />
    <tag k="" v="" />
  </way>
  <relation id="">
    <member type="" ref="" role="" />
    <tag k="" v=""/>
  </relation>
</osm>
```

Table 3. A comparison of the OSM model and a traditional model

Geodata model	Geometry and topology	Geometry validity	Layers	Attribute fixing
Traditional GIS (i.e. ESRI Shapefile)	point polyline polygon	required	by theme and by geometry	by layer
XML OSM	point polyline relation (a group of primitives with roles)	low-level	all in one	no

The export to other formats is nontrivial and ambiguous (See Table 3). In geometry and topology, typical problems occur, which concern multipolygons, roles of elements, aggregations, routing restrictions and routes. A chronic complication is the flow design of attributes, their immediate in ad-hoc changes by users [39]. The improvement of present data exchange between OSM and GIS software depends perhaps on the GIS developers, because big GIS stations are not typical tools in the community [26].

Real usage in GIS:

- Raster map over slippy maps (i.e. Quantum GIS with plugin, ArcGIS 10 default)
- Raster map via WMS [@1]
- ESRI Shapefiles ready to download [@2,@3] (i.e. layer of road, railway, waterway, buildings, natural and POI, without the information in *relations*)
- PostgreSQL/Postgis (conversion by script Osm2pgsql) [23]

Actually on-line accessible OSM exports to ESRI Shapefiles suffer from these losses:

- big polygons aggregated from multiple parts in relation
- rings composed of polygons
- routes, i.e. cycle, hiking, public transport
- parts of attributes
- many themes, i.e. landuse, traffic accessories, routings data
- many features stored in atypical objects (i.e. a POI saved in a polygon of building, instead of a point)
- areas of administrative terrirorial units are exported only as polylines of borders

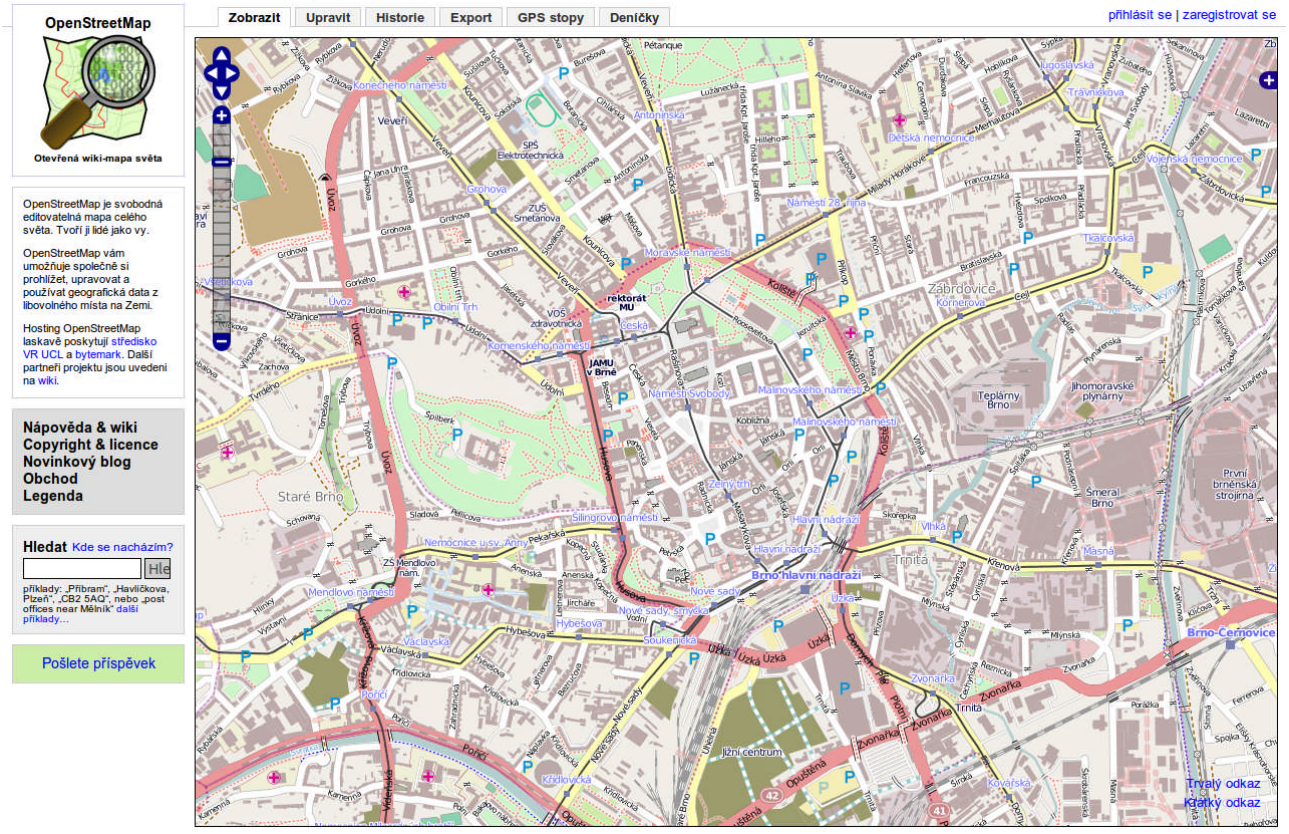

Fig. 2. Detail of the centre of Brno (Mapnik renderer) [32]

For Quantum and ArcGIS, there exist trial plugins for direct editing of OSM data, but their installation and operation is not trivial and have low support in the community. [24,25]

For the web, for print (PNG, SVG) and mobile devices (typically cell phone or GPS navigation) exist special conversional software or online pre-prepared data [28].

For the import of free data into OSM there are many special scripts and methodology. A good example is a versatile import utility ogr2osm written in Python, which uses the potential of formats that can be read by the linked library GDAL/OGR. However, reading the OSM format and exporting via GDAL/OGR to other geodata formats is not supported. [27].

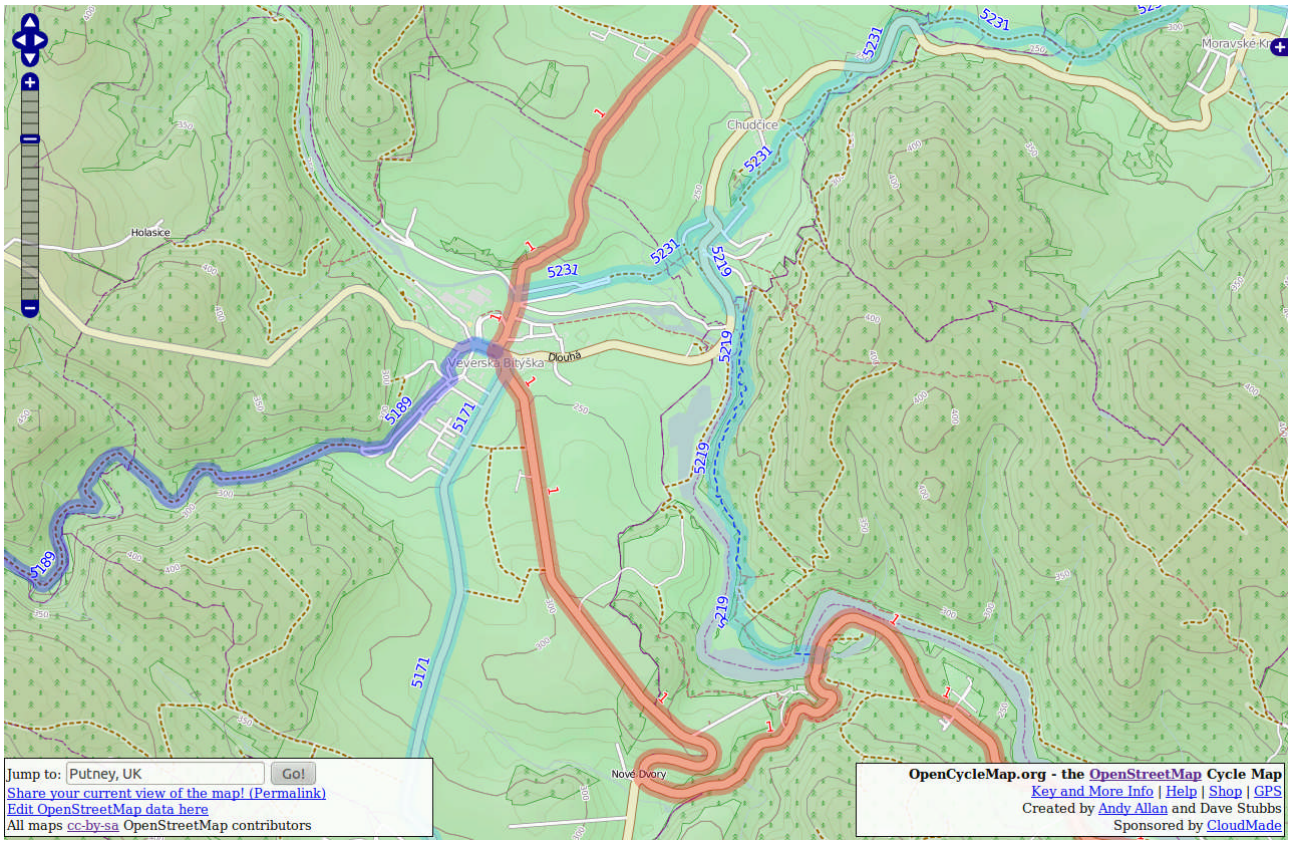

Fig. 3. Detail of cycle route near Brno reservoir (OpenCycleMap renderer) [33]

MAPPING STATE IN THE CZECH REPUBLIC

In every country the state of the map is different, mostly affected by available free datasets for import or creating geodata [5].

Top data themes

- *Complete* country and regional highways – motorways, expressways and the 1., 2., 3. classes of roads (all in the real course of way). The second class of highways was partially imported from corporation HS-RS. *In progress* - forest tracks and field tracks, village and town residential ways, accessories, options...
- *Complete* railway - main and regional railway with stations. *In progress* - spur, yard, siding track, attributes (electrified, voltage system)
- *Complete* administrative system borders – derived from cadastral territories (CUZK KM), from villages to regions, consecutively the whole republic according to NUTS and LAU standards. *In progress* – fixing the hierarchy.
- *Complete* waterway - imported datasets DIBAVOD (streams, rivers, lakes, basins, marsh). *In progress* – checking and fixing duplicates between old data and import
- *Complete* forest – imported from UHUL layer of forests derived from PUPFL (cadastral parcels primary designated for forest). *In progress* – some fixing of true sizes of forest

- Cca 10% of house addresses – the free part of addresses from UIR-ADR (old version from the ministry MPSV, only former district towns). *In progress* - next 80 % will be imported semi-automatically [29]

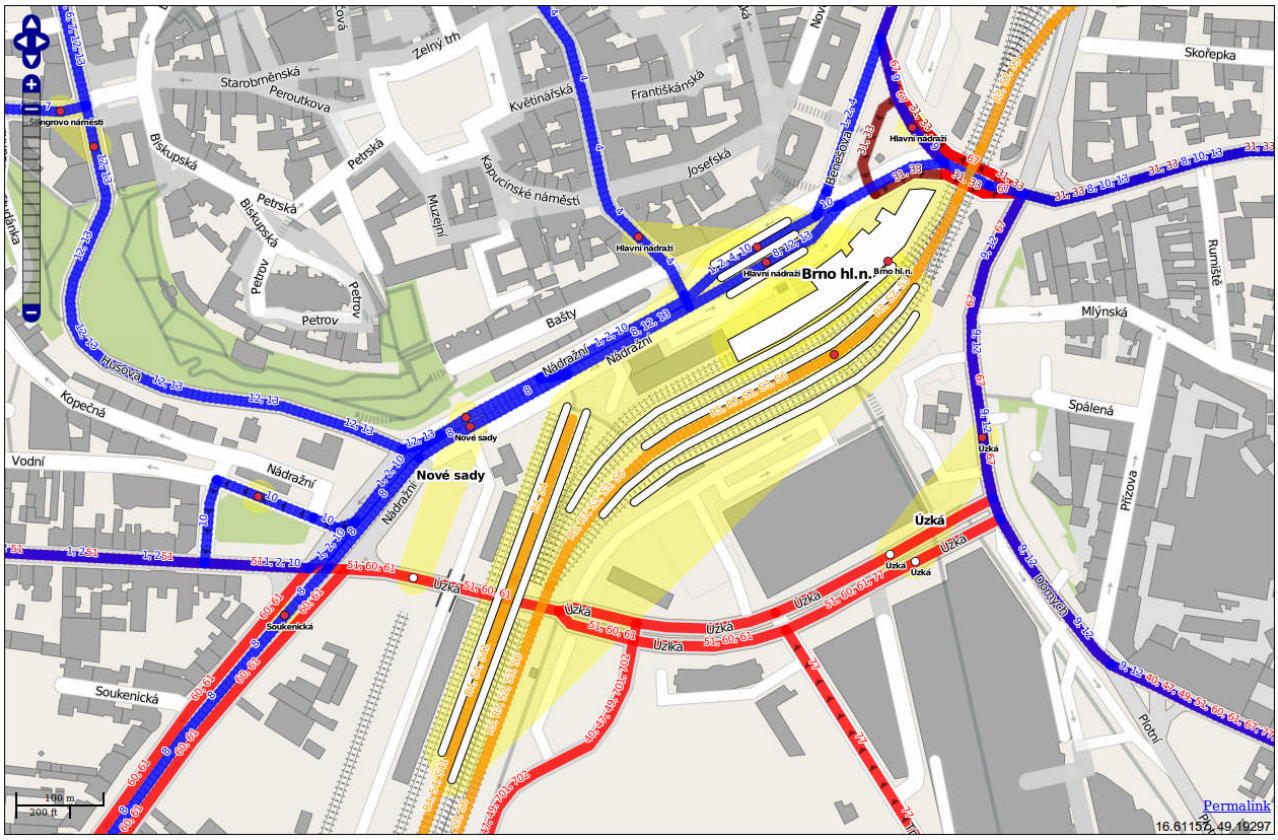

Fig. 4. Detail of public transport map for Brno main station (Öpnvkarte renderer) [34]

Other theme in progress

- Buildings (semi-automatic)
- Hiking, cycle, mountain bike and ski ways
- Public traffic services
- Routing limits
- Landuse/landcover
- High voltage power infrastructure
- Points of interests

Data sources for visual interpretation

- Cadastral map – WMS ČÚZK
- Aerial orthophoto map – WMS UHUL (captured before 2000), spatial resolution 1 meter
- GPS tracks – users' self-made
- Series of street photo – NORC (only Prague, Brno, Ostrava, Pilsen, Olomouc)
- Satellite orthophoto map – Microsoft Bing (only 16 % of the area of the Czech Republic – i.e. Prague, Brno and surroundings), spatial resolution cca 0,5-1 meter; map composed from satellite images from the company GeoEye, acquired mostly in 2003 or 2007.

- Satellite orthophoto map – Yahoo (only 3 % of the area of the Czech Republic - Prague and surroundings), spatial resolution cca 1 meter, images acquired from the satellite Ikonos in 2007

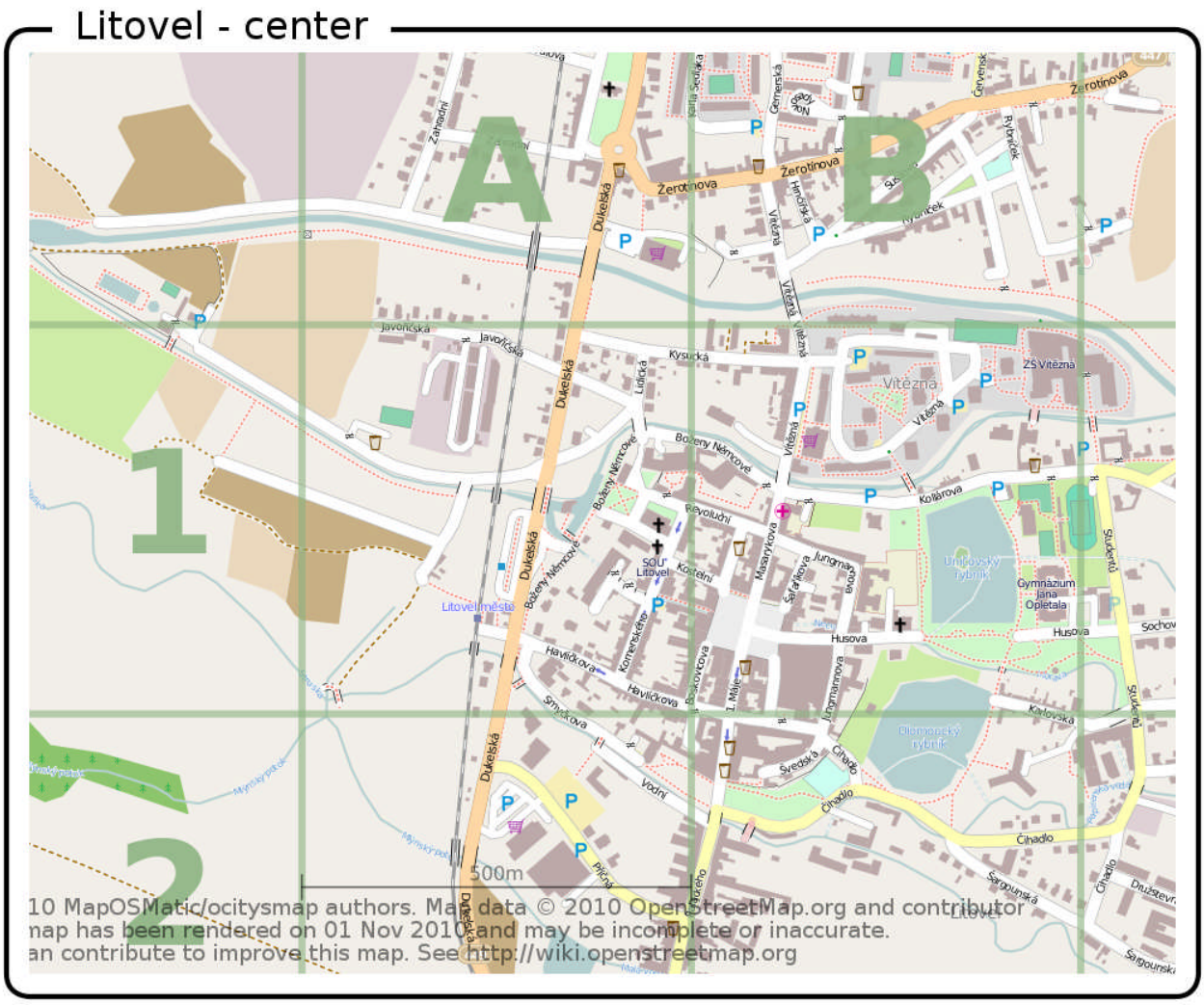

Fig. 5. Detail of the center of Litovel town prepared for print (MapOSMatic via Mapnik renderer) [35]

THEMATIC MAP

For active users' motivation a visual demonstration of their work is very important, because it enables to see their own tile in the mosaics and it brings inspiration for the others. For the other consumers it demonstrates the summary of the database. A typical generalized way from geodata to a webmap comprises OSM XML format > database (PostreSQL) > renderer (Mapnik/Osmarender) > www server + Openlayers [40, 41].

Lists of online map service, that affect the area of the Czech Republic[37]:

- Base, street, road map's (See Figure 1, 2 and 6) [@4-@7]:
- Foot, cycle (see Figure 3), hiking, mountain bike, orienteering and wheelchairs map's [@8-@14]
- Power line and grid map [@15]
- Public transport service map (see Figure 4) [@16]
- Road navigation [@17, @18]
- Map with ship navigation feature [@21]

- Isometric map [@19]
- Paper map (see Figure 5) [@20]
- Meta-information map service [@22, @23]
- Map compare [@24]
- Map overlay [@25]

CONCLUSION AND FUTURE

In the present and in the near future a great change of the license is going on. During the next year, secondary map services like routing should be improved. A great theme is now how to match a stabilized attribute system with quality standards and validation, but still open or under the control of the community [14].

In the present moment, geodata and maps from OSM - compared to other map sources on the Czech market – are interesting foremost by their price, accessibility and actuality. In many geographic themes, the area completeness in general is ready-to-wear in many places (i.e. towns). Nevertheless, a great part of he Czech Republic has a low level of mapping quality.
For a GIS user, OSM can be a valuable help, a secondary or other view of the same place, and may support or complete GIS analyses. For ordinary people OSM represents an effective tool to get and share geographic knowledge in a legal way.

NOTES

[@1] http://www.osm-wms.de/

[@2] http://download.geofabrik.de/osm/europe/

[@3] http://downloads.cloudmade.com/europe/czech_republic

[@4] http://www.openstreetmap.org/

[@5] http://www.openstreetbrowser.org/

[@6] http://www.mapsurfer.net/

[@7] http://open.mapquest.co.uk/

[@8] http://tchor.fi.muni.cz:8080/

[@9] http://hikebikemap.de/

[@10] http://openstreetmap.cz/

[@11] http://opentrackmap.no-ip.org/

[@12] http://gismaster.liberix.cz/osmoverlay.html

[@13] http://www.openpistemap.org

[@14] http://wheelmap.org/

[@15] http://energy.freelayer.net/

[@16] http://openbusmap.org/

[@17] http://www.openrouteservice.org

[@18] http://www.yournavigation.org/

[@19] http://osm.kyblsoft.cz/3dmapa/

[@20] http://maposmatic.org/

[@21] http://www.freietonne.de/seekarte/

[@22] http://www.qualitystreetmap.org/osmqa/

[@23] http://openstreetbugs.schokokeks.org/

[@24] http://tools.geofabrik.de/mc/index.html

[@25] http://sautter.com/map/

REFERENCES

[1] Clinton, S. (2009) Crowdsourcing is radically changing the data landscape: Case study of OpenStreetMap, 24th International Cartographic Conference, 15-21th November 2009, online: http://icaci.org/documents/ICC_proceedings/ICC2009/html/nonref/22_6.pdf

[2] Bárta, D. (2008) Projekt OpenStreetMap z pohledu geoinformatika, Geoinformatics FCE CTU 2008, 18.-19. September 2008, pp. 91-108, online: http://geoinformatics.fsv.cvut.cz/pdf/geoinformatics-fce-ctu-2008-03.pdf

[3] Rapant, P.: Geoinformatika a geoinformační technologie. VŠB-TU Ostrava, 2006. 500 str. ISBN 80-248-1264-9, pp. 304. http://gis.vsb.cz/rapant/publikace/knihy/GI_GIT.pdf

[4] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/FAQ#Why_are_you_making_OpenStreetMap.3F

[5] OpenStreetMap, wiki pages, accessed 5. December 2010, http://wiki.openstreetmap.org/wiki/WikiProject_Czech_Republic/freemap

[6] Wikipedia, the free encyclopedia, GNU_General_Public_License, GRASS_GIS, Linux, PostgreSQL, Mozilla, SETI@home, OpenOffice.org, PostGIS, Wikipedia, Quantum_GIS, OpenStreetMap, accessed 30. October 2010, http://en.wikipedia.org/

[7] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Beginners_Guide_1.3

[8] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/API_v0.6#Elements_2

[9] OpenStreetMap, wiki pages, accessed 30. October 2010, http://planet.openstreetmap.org/

[10] OpenStreetMap, stats pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Planet.osm

[11] OpenStreetMap, wiki pages, accessed 30. October 2010, http://munin.openstreetmap.org/openstreetmap/smaug.openstreetmap/postgres_size_openstreetmap.html

[12] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/OpenStreetMap_License

[13] OpenStreetMap, wiki pages, accessed 30. October 2010, http://www.osmfoundation.org/wiki/Main_Page

[14] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/State_Of_The_Map

[15] OpenStreetMap, talk-cz archive pages, accessed 30. October 2010, http://www.mail-archive.com/talk-cz@openstreetmap.org/msg05938.html

[16] Gale, G (2010) What's Wrong With OpenStreetMap? Have Your Say, accessed 30. October 2010, http://www.vicchi.org/2010/07/07/whats-wrong-with-openstreetmap-have-your-say/

[17] Chance, T (2010) Political philosophy in Open Street Map, accessed 30. October 2010, http://tom.acrewoods.net/2010/07/16/political-philosophy-in-openstreetmap/

[18] Haklay, M (2010) The tyranny of place and OpenStreetMap, State Of The Map 2010, 9-11th July 2010, http://povesham.wordpress.com/2010/07/10/the-tyranny-of-place-and-openstreetmap/

[19] OpenStreetMap, tool pages, accessed 30. October 2010, http://tagwatch.stoecker.eu/

[20] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Stats

[21] Janda, A.: Aktivita uživatelů OpenStreetMap v ČR, accessed 30. October 2010, http://osm.kyblsoft.cz/statistiky/aktivita-uzivatelu

[22] Budhathoki, N.R., Haklay, M., Nedovic-Budic, Z. (2010) Who are the mappers and why do they map in OpenStreetMap? State Of The Map 2010, 9-11th July 2010, http://www.slideshare.net/nbudhat2/nama?from=ss_embed

[23] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Osm2pgsql

[24] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/QGIS_OSM_Plugin

[25] ESRI: ArcGIS Editor for OpenStreetMap, accessed 30. October 2010, http://esriosmeditor.codeplex.com/

[26] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Converting_map_data_between_formats

[27] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Ogr2osm

[28] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Converting_GPS_track_data_between_formats

[29] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/Import_Adres_%C4%8CR

[30] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/State_Of_The_Map_2010

[31] OpenStreetMap, map, accessed 30. October 2010, http://www.informationfreeway.org/?lat=49.7957&lon=15.4092&zoom=8&layers=0FB0F0

[32] OpenStreetMap, map, accessed 30. October 2010, http://www.openstreetmap.org/?lat=49.19448&lon=16.6092&zoom=15&layers=M

[32] Moos, M. (2010) Öpnvkarte, accessed 30. October 2010, http://openbusmap.org/?zoom=17&lat=49.19012&lon=16.61099&layers=BT

[33] Allan, A. (2010) OpenCycleMap, accessed 30. October 2010, http://www.opencyclemap.org/?lat=49.2718&lon=16.4494&zoom=14

[34] Petazzoni, T., Petazzoni, M. : MapOSMatic, accessed 30. October 2010, http://maposmatic.org/jobs/20197/IKAojBnIYIClSfsZ#submitmapform

[36] Freemap SK, wiki pages, accessed 30. October 2010, http://wiki.freemap.sk/VojenskeMapyCSLA

[37] OpenStreetMap, wiki pages, accessed 30. October 2010, http://wiki.openstreetmap.org/wiki/List_of_OSM_based_Services

[38] OpenStreetMap, map, accessed 5. December 2010, http://www.informationfreeway.org/?lat=49.98521&lon=14.39959&zoom=16&layers=0FB0F0

[39] OpenStreetMap, wiki pages, accessed 5. December 2010, http://wiki.openstreetmap.org/wiki/Map_Features

[40] OpenStreetMap, wiki pages, accessed 5. December 2010, http://wiki.openstreetmap.org/wiki/Slippy_Map

[41] OpenStreetMap, wiki pages, accessed 5. December 2010, http://wiki.openstreetmap.org/wiki/Mapnik