The Commercialisation of Public Data - How Does Participatory Data-mining Look on a Global Scale?

Jiří Pánek

Department of Development Studies, Palacky University in Olomouc, Czech Republic, and Department of Geography, University of South Africa, jirkapanek@gmail.com

Abstract

The commercialisation of local spatial knowledge is an emerging problem across communities all around the world. Users creating spatial and attributive data through the use of commercial tools such as Google Map Maker, are extending corporate databases, but aren’t able to use the very same data freely when they need to. Open-source and participatory mapping projects such as OpenStreetMap are designed to support community empowerment and to ensure that high quality data are available to everyone regardless of their origin, social status and position within the power structure. The article points out the danger of commercialisation of public data using the example of community development projects. It also shows the advantages of participatory data-mining, using as an example the OpenStreetMap project in the Kibera slum, Nairobi Kenya and shows the possibilities of such an initiative in Koffiekraal, NW province of South Africa.

Keywords
Participatory GIS, PGIS, OpenStreetMap, Google Map Maker, local spatial knowledge (LSK), Koffiekraal, Kibera

1. Introduction

Community mapping as an independent methodology historically comes not only from Participatory Rural Appraisal (PRA), but from all other visual methods of PRA, like matrix scoring, seasonal diagramming, Venn diagramming, ground mapping, it has been most widespread. This is mainly due to the versatility and power of community mapping, the relative ease with which it can be facilitated and the fun, fulfilment and pride, which people derive from it. Nevertheless, local spatial knowledge (LSK) transformed into community maps has already been used by Non Chi Ning Ga, an Iowa Indian chief, as part of a land claim presented at a U.S. government-hosted council designed to persuade several midwestern tribes to agree to land cessions and new treaties in Washington, in 1837 (Chambers, 2006; Turnbull & Watson, 1993).

Many critics of GIS were concerned that the technology was being employed with the explicit goal of expanding political and economic control over those already disadvantaged by local, regional, and global divisions of power. They were therefore afraid that GIS was still only a tool of control and technological surveillance. Obermeyer (1998) states that GIS as a tool of technological control was mainly developed and practised by white males employed in academic and governmental institutions in North America and Europe. There was growing concern among
academia at that time, that GIS may become a tool of techno-neocolonialism. In the mid-1990s there were pioneering attempts to devise alternative approaches, notably through work by Abbot et al., 1998; Goodchild, 1991; Harris & Weiner, 1996, 1998; Openshaw, 1991; Weiner, Warner, Harris, & Levin, 1995.

All these critics believed that, within most communities, mapping was historically understood as something that would misuse local spatial knowledge for external benefits and bring no good for the community empowerment.

With the rise of GIS in general the development workers stared exploring the subject of using georeferenced mapping within their practices as well. Participatory GIS (PGIS) emerged from participatory approaches to planning and spatial information and communication management. As a term, PGIS came to express the adoption of GIS in order to empower indigenous and local communities in their daily lives. It represents the vision of GIS practitioners who have developed an interest in the socio-political contributions and implications of the technology and its ability to empower less privileged groups in society. (Abbot et al., 1998; Pickles, 1995).

This article discusses the issue of global PGIS as it currently stands, using the examples of Google Map Maker and OpenStreetMap – two major mapping services with worldwide coverage. This article critically answered the following questions:
(i) How can one practice community mapping in the rural village of South Africa?
(ii) How should one choose a mapping method and platform for the project?
(iii) How can the Koffiekraal community use data gathered on the OpenStreetMap for its own development?

2. Historical and Spatial Context

As the colonial history of mapping in Africa was originally based on three main needs – business, administration and war, it was quite visible that maps were another tool of control and technological surveillance. Commercial mapping was done mainly in German South-West Africa by Deutsche Kolonialgesellschaft für Südwestafrika and Deutsche Diamanten-Gesellschaft. Great Britain mapped its colonies mainly in order to facilitate the process of administration and tax collection. In the African colonies of Great Britain, the Indian model, which worked well under control of the Survey of India (SoI), was used. France used le Service Géographique de l’Armée (SGA) and as the name suggests it was mainly military mapping that was conducted by this organisation. (Collier, 2005; Moser, 2005, 2006).

The use of participatory approaches in GIS and in mapping emerged from the critics of GIS in the 1990s. The critics were afraid of classical GIS and mapping methods, mainly because of the need for highly trained experts and the high cost of software, hardware and data. Based on this fear community mapping brings the ability to use GIS in order to assist communities which are directly
linked to the developing program and agendas. Community mapping is part of a more socially aware type of GIS which gives greater privilege and legitimacy to local or indigenous spatial knowledge. Variously labelled as, Participatory GIS, Public Participation GIS, and Community-integrated GIS, these newer approaches are context- and issue-driven rather than technology-led and seek to emphasize community involvement in the production and/or use of geographical information. These approaches seek the human story behind the data and allow classical development workers to start using GIS within their work. Recent years have witnessed a burgeoning of applications of GIS, which grant legitimacy to local spatial knowledge. By incorporating various forms of community participation these newer framings of Geographical Information Systems are response to the critiques of GIS, which were prevalent in the 1990s. The main argument of “democratisation of GIS” is used to justify the legitimacy of PGIS and similar approaches inside the GIS context (Abbot et al., 1998; Dunn, 2007; Pánek, 2011).

The socio-economic benefits of GIS have been criticised by different scholars for under representing the people excluded from the society (Harris & Weiner, 1996; Sheppard, 1995), for social implications of GIS as a technology (Pickles, 1995; Smith, 1992), for increasing the surveillance upon the society (Curry, 1995; Goss, 1995; Pickles, 1993), for being a black-box technology (Curry, 1995; Goss, 1995) and for being just a profit motivated technology (Veregin, 1995). In 2000 Nadine Schuurmann (2000) introduced the term Critical GIS and her discussion shed new light on the ways in which PGIS can have further impact on the exploitation of local spatial knowledge. Critical to this widening participation, however, is the need for PGIS scholars and practitioners to be more explicit about who “the public” is and what “participation” means if appropriate goals are to be achieved (Schlossberg & Shuford, 2005).

In this paper the author understands the “public” as a group of users, or a community, who are actively participating in creating and using spatial and attributive data, while that process of creating data he is referring to as “participation”. He believes that the process of communities creating their own data has an empowering effect, because members of the community (“public”) have the opportunity to think spatially about their environment and literally put their community on the map. The process of creating the data triggers feelings of belonging to the community and the ownership of the process. With ownership empowerment starts and leads to sustainable development, driven and run by the community itself. This has been observed during the community mapping in Koffiekraal and other villages in the Bojanala region (Vlok & Pánek, 2012). One can always argue that this is a simplistic vision of the process and that there is a need of much deeper assessment of this statement and one would not be far from truth. On the other side Chambers (2006) as well as Amsden & VanWynsberghe (2005) or Corbett et al. (2006) mention how important it is for communities to be the driven power in the mapping process and how much of empowerment and involvement one can derive from it.
3. Crowdsourcing Within the PGIS Discourse

The word Crowdsourcing is a compound contraction of two words - Crowd and Outsourcing. Thus Crowdsourcing means outsourcing to the crowd. The origin of this word itself is unknown, which is typical of the Web 2.0 phenomenon – an anonymous user launched the term for the first time on an Internet Forum. The term was popularised by Howe in an article “The Rise of Crowdsourcing”, which was published in Wired (Howe, 2006).

Howe (2006) proposed the following definition:

Simply defined, Crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by individuals. The crucial prerequisite is the use of the open call format and a wide network of potential labourers.

From a geographical point of view, Crowdsourcing is used largely in several projects, such as OpenStreetMap or Google Map Maker. OpenStreetMap creates and provides free geographic data such as street maps, points of interest and areas of interest. Some of the data are original, which means they were created by OpenStreetMap users. On the other hand parts of the data were adopted from various sources. This adoption of data should, according to the OpenStreetMap Licence policy, follow strict rules. The project was started in 2004 by Coast as a collaborative project, to provide users all around the world with free spatial data. Coast believed that most map users thought of current data as free, but it actually has legal or technical restrictions on its use, which hold users back and prevents their creativity with spatial data. His original idea was to enable users to be creative with data and share it freely. There is a great deal of controversy about OpenStreetMap data and its origin. Nevertheless, OpenStreetMap data are, in some cases, much more detailed than data provided by major commercial map services such as Google Maps, Yahoo Maps and Bing Maps (OpenStreetMap Wiki, 2013).

The following comparison (see Figure 1) shows the Kibera slum in Nairobi, Kenya on OpenStreetMap and Google Maps, where intensive mapping was organised within the Map Kibera and the Map Kibera project (Map Kibera Project, 2013; Map Kibera, 2013). The Map Kibera Project aimed at producing reliable and clear data and maps using GIS (mainly the freeware Quantum GIS), and making them available on the Map Kibera Project’s official website, while Map Kibera is a crowdsourcing initiative started in November 2009 when Kibera was a blank spot on the OpenStreetMap and on Google Maps. Young Kiberans created the first free and open digital map of their own community. Map Kibera has now grown into a complete interactive community information project, including local radio and online news. Kibera is one of the largest slums in Africa, situated in Nairobi, Kenya. Many UN agencies, including UN-HABITAT, US Government agencies such as USAID, and NGOs such as Carolina for Kibera, have a presence nearby in...
Nairobi, and as a result Kibera is one of the most well-known, researched, and serviced slums in the world. Despite this focus, Kibera was literally a blank spot on the map, its patterns of traffic, scarce water resources, limited medical facilities, etc. remained invisible to the outside world, and to the residents themselves. Without a basic knowledge of the geography of Kibera it is impossible to have an informed discussion on how to improve the lives of residents of Kibera. Members of the Map Kibera initiative have produced the first complete free and open map of Kibera. In November 2009 local young people learned to create maps using OpenStreetMap techniques. This included surveying with GPS, and the digitisation of satellite imagery and paper based annotation with Walking Papers. Data consumers were consulted for their needs, to help add direction to the feature types collected, and aided in immediately making use of the map data (Hagen, 2011; Map Kibera, 2013; Marras, 2009).

Currently the Map Kibera Project (2013) allows users to download data and maps only for the Kinda village, which is the most western part of Kibera. These data include information about terrain, structures, sewage system, water points, electricity supply, public toilets & lavatories, building materials, business distribution or population. For the whole Kibera only the terrain and structures are available. Unfortunately one can only view the data in pdf, jpg or kmz. If desired in shp format one needs to contact the mapping team with the data access request. On the other hand the Map Kibera (2013), based on the OpenStreetMap, allows users to download shapefiles with information about the transport, health, security, education, religion, boundaries and polling places. Furthermore the additional CSV data about health and education are available too.

There is a fundamental difference between the OpenStreetMap and Google Maps of the Kibera slum and this difference in not just in the data coverage, but also in the data structure. While Google Maps use for the Kibera area only single style roads, there are at least three types of roads on the same zoom level for the OpenStreetMap. Regarding the accuracy at the OpenStreetMap, roads are mostly digitalised based on the Bing aerial photographs or the GeoEye imagery with high resolution (50 cm), while points are gathered by Map Kibera crew (Map Kibera, 2010).

The points of interest (POI) mapping was realized in more detail in four issue areas: Health, Security, Education, and Water/Sanitation. Some of the mappers also carried digital cameras or Flip camcorders.
Both OpenStreetMap and Google Maps data are worldwide and their crowdsourcing activities are visible elsewhere, not just in Kibera. After the earthquake in Haiti, on January 12, 2010, both services tried to improve their user-generated map coverage of Haiti, unfortunately there was a duplication of effort as well as barriers to combining data sets generated within different software packages. This issue is best illustrated via the lack of compatibility between OpenStreetMap and Google’s Map Maker. Following the earthquake people utilised both services and started to trace roads, hospitals, and other sites of interest. Unfortunately, due to licensing issues (OpenStreetMap issues all map data using a Creative Commons license, but Google retains the intellectual property of all information created using Map Maker), data is not portable between the two systems and efforts were undoubtedly duplicated. More importantly, this incompatibility resulted in maps with varying degrees of coverage, depending upon the location within Haiti. Figure 2 demonstrates that different parts of Haiti had varying levels of coverage in OpenStreetMap (blue shading) and Map Maker (yellow shading) (Po Ve Sham – Muki Haklay’s personal blog, 2010a). Figure 2 shows for each grid square of 1 km, which of the datasets contains more information in terms of roads length. The file contains the total roads length for both datasets and then it is calculated difference between them using the equation: \( \sum (\text{OSM roads length}) - \sum (\text{Map Maker roads length}) \) for each grid square (Po Ve Sham – Muki Haklay’s personal blog, 2010a). Google Map Maker is covering 354 km\(^2\) which are not covered by OpenStreetMap, and OpenStreetMap is covering 1044 km\(^2\) that are missing from Google Map Maker, so clearly there is a benefit in integrating them (Po Ve Sham – Muki Haklay’s personal blog, 2010b).
While the varying levels of information are not debilitating, this example illustrates the challenges that integrating crowdsourced data can pose. Not only should one be concerned about quality and the ground-truth, but issues of intellectual property and regulation can complicate such collaborative efforts. Despite these concerns, OpenStreetMap and Map Maker did continue to provide rescue efforts with street map coverage extremely quickly. These spatial data were ultimately crucial for first responders, aid workers, and even U.S. military humanitarian efforts on the ground (Zook, Graham, Shelton, & Gorman, 2010).

4. The Agreement Between the World Bank and Google

On 17th January 2012 the World Bank (WB) announced an agreement with Google Inc. aimed at improving the ability of developing countries to access a web-based community mapping tool and data to help in better monitoring public services, and improve disaster and humanitarian response efforts. Most developing countries do not have basic local data about where schools, hospitals, and water points are located, and the data they have is often outdated or incorrect. One way to collect this information is by asking citizens directly and crowdsourcing the locations of public infrastructure. Under this agreement, Google will provide the World Bank and its partner organisations - including governments and UN agencies - with access to Google Map Maker’s underlying geospatial data, which includes detailed maps of more than 150 countries. Through this tool, citizens are able to directly participate in the creation of maps by contributing their local
knowledge, and those additions are then reflected in Google Maps and Google Earth. These maps include locations such as schools, hospitals, roads and water points, that are critical for relief workers to know about in times of crisis; and will help NGOs, researchers, and individual citizens to effectively identify areas that might be in need of assistance (World Bank, 2012).

The idea is similar to the one, which is fundamental for the OpenStreetMap project. The difference is the issue regarding the copyright and property rights related to the data obtained through the mapping process. Once the local spatial data are crowdsourced and saved to the Google Maps database, local communities are not able to use them in any other way except via the Google Maps services. This process can be described as “commercialisation of public data”.

The agreement basically restricts users from using data with any tools that Google Inc. does not approve of. It is understandable that global companies, such as Google, are aiming to increase their positions in the current digital world and data crowdsourcing is one way of achieving this goal. Surprisingly the World Bank has joined this venture, despite its previous projects supporting the Open Data community, e.g. Kenya OpenData (2013) project, funded by the World Bank in 2011.

After the World Bank announced the agreement, various bloggers and NGOs (Kelso’s Corner, 2012; Owen abroad, 2012; TechPresident, 2012) criticised WB for this move and suggesting alternative approaches, such as Ushahidi or OpenStreetMap. As a result of a number of critiques, the World Bank distanced itself from closed Google Maps Data (ReadWrite, 2012) clarified its open data policy with the statement “...the World Bank only supports citizen-mapping efforts that give users free access to the map data they create. and ... our agreement does not extend to supporting new citizen mapping activities or data creation through the Google Map Maker platform. Therefore, when it comes to creating map data, we will work with a variety of other mapping platforms consistent with our development objectives, terms of use and guiding principles around open data.” (Inside the Web, 2012).

5. Case study of Koffiekraal

In the Local Agenda 21 Planning Guide (developed out of the United Nations Rio Conference on the Environment in 1992), community-based mapping is identified as a best practice for locally-based sustainability planning (IDRC, 1996). Based on the example of the Kibera slum, in combination with Community Assets Mapping Methodology (CAMP), the community mapping was executed in Koffiekraal, North-West province, South Africa.

Although conceptualised independently, the CAMP methodology shows many similarities to the well-known methodology of ABCD (Asset Based Community Development) developed in the 1990’s by Kretzmann and McKnight (1993) in the USA. As the name suggests, mapping is part and parcel of the CAMP methodology to effect change in communities. These mapping activities include power mapping, association mapping, economic mapping, cognitive mapping and creative
mapping — collectively aimed at the communities being able to visualise their community assets, capacities, abilities and community structures.

Besides the CAMP methodology, the field mapping and transformation of data into the OpenStreetMap was undertaken. The community members were active participants in the whole process and their needs, such as water points mapping, were taken into the consideration.

The results from the mapping were either digitalised, when needed, or directly uploaded to the OpenStreetMap. This action created the most accurate and spatially rich map of the Koffiekraal currently available. The difference between the coverage of Koffiekraal before (Fig. 3) and after (Fig. 4) the mapping programme is easily visible. In total there were almost 90 point features and 10 polygon features added to the OpenStreetMap. Additionally dense road network consisting of 1174 road elements, and river were also digitalised based on the aerial photos delivered by the National Geo-Spatial Information (NGI) South Africa. The accuracy of the mapping was originally around 5 metres, as the coordinates were collected by handheld GPS device Garmin 60CSx. During the process of the digitalisation and import to the OpenStreetMap, points´ and polygons´ coordinates were adjusted according to the NGI aerial photograph.

The ability of OSM to tag points, lines and polygon was used especially while creating POIs and roads network, where mostly tags: highway=unclassified, highway=residential, highway=track or highway=footway were used. Roads were tagged based on the in-field observation and assessment. There can be small inconsistency in these data as some of the tagging was done as a part of the post-processing in the office. The mapping efforts in Koffiekraal did not end in November 2012, but are running continuously with the aim to make Koffiekraal the showcase of participatory mapping with use of OpenStreetMap in that region.

It is believed that communities creating their own maps can experience empowerment through the process. They have the opportunity to think spatially about their environment and literally put their community on the map. The process of creating data and eventually a map, triggers feelings of belonging to the community and a sense of ownership of the process. With ownership empowerment starts and leads to sustainable development — driven and run by the community itself (Vlok & Pánek, 2012).
Figure 3: Difference in coverage of Koffiekraal between OpenStreetMap and Map Maker
(21st August 2012)

Figure 4: Difference in coverage of Koffiekraal between OpenStreetMap and Map Maker
(2nd November 2012)
6. Advantages of the OpenStreetMap for Community Projects

The OpenStreetMap community is creating open data all around the world from Belarus to the Gaza Strip. Thus it is fundamental that the independent position of community initiatives is supported and strengthened. The agreement of Google Inc. and the World Bank can set a precedent that will affect community-based mapping projects all over the world. This is the first case on such a scale when a private mapping company creates an agreement with an organisation representing different state members on the provision of spatial data, which are, on a certain scale, collected via crowdsourcing methods.

Erle (personal communication, 2012) asks how does this support global economic development? “The only economic development happening through Map Maker is Google's. By encouraging their beneficiaries to use it, the World Bank is blindly leading them into nothing more than digital serfdom.” Similar concerns were already mentioned by Pickles (1995) and the main question, almost two decades later, is whether the international open data community can learn from its previous mistakes.

Despite the agreement of the World Bank and Google Inc., there is a growing adoption of OpenStreetMap services. OpenStreetMap gained two new big services in March 2012 with the disclosure by FourSquare and Apple that they have chosen the crowdsourcing map as background data for their products. Although Apple is using its own version of maps based on TomTom and labels in iPhoto from OpenStreetMap. This news also marked a departure from Google Maps as a mapping source for these two companies. Open MapQuest (2013) is currently also testing the option of utilizing OpenStreetMap data for their mapping service.

The current market with online mapping services is biased towards commercial companies, represented by Google Maps, Bing Maps and Yahoo Maps among others and mapping services based on the OpenStreetMap database. Community projects can benefit from this bias as both groups are currently trying to convince the general public about their better coverage and other advantages.

OpenStreetMap has many positive and negative characteristics in terms of providing a comprehensive and accurate mapping resource. The coverage and accuracy of OpenStreetMap is loosely connected with the number of volunteers mapping a given area and the position of the mapping locations. As stated in Haklay (2010) there is evidence to suggest that there are “areas where nobody wants to map”. If this is a widespread problem in OpenStreetMap it represents a significant obstacle to improving accuracy and coverage. This is an area where Google Maps and Bing Maps, as well as other commercial mapping services, have a distinct advantage (Ciepluch, Jacob, Mooney, & Winstanley, 2010).
Based on the examples in this paper I am confident to state that there are advantages to OpenStreetMap for community mapping projects. The advantages can be seen in open and free access to the data created by the community, as well as licensing protection of the data, which means that OpenStreetMap data cannot be sold or used for commercial purposes due to the Creative Commons licensing. The editing features and vast amount of supporting software makes OpenStreetMap a very powerful tool for community mapping, and the worldwide community of users is an example of a crowdsourced library of experience and knowledge related to the tool. The export feature allows users to download any data from the OpenStreetMap and use it as a starting point for real application of the Participatory GIS. This ability was used for Koffiekraal water points service areas analysis by Pánek (2013). In that paper the author shows the link between the Koffiekraal participatory mapping and Local Spatial Knowledge regarding the waterpoints and presents the usage of the GIS as a tool for informed community decision about the possible locations for new waterpoints.

7. Conclusions

This article has presented the different approaches of Google Maps and OpenStreetMap, as examples of participatory data-mining services which use crowdsourcing to gather data from the community of users worldwide. It also gives a brief summary of the agreement between Google Inc. and the World Bank, which has provided the World Bank with access to Google Map Maker’s underlying geospatial data, including detailed maps of more than 150 countries. Lastly, this article discusses the example of good practice in Kibera, Kenya which could be implemented in Koffiekraal, South Africa. The Koffiekraal experience provided an example of how a community can pool its efforts to do more extensive mapping and thus literally and figuratively put themselves on the map.

Local communities need to be aware of the possibilities and implications of using one of the presented participatory approaches to digital mapping. These approaches are either community-driven mapping processes, as represented by the OpenStreetMap project, or corporate-driven mapping processes, as represented by the Google Map Maker project. Both of these processes have advantages and disadvantages, and similarly they have individuals who support them and those who are against their use.

The decision about which service to use should be taken at the beginning of the whole participatory mapping process and the local community should have the detailed SWOT analysis of both options with respect to local realities. Using GIS as a tool for community building and community empowerment is an emerging practice. There were evidences provided for using OpenStreetMap as a mapping platform for community projects and author would like to invite other practitioners to share their experiences with different mapping services and the challenges they face during the participatory mapping processes.
8. Acknowledgement
The author would like to thank and acknowledge the financial and logistical support provided by the Greater Rustenburg Community Foundation (GRCF) and the University of South Africa (UNISA) Department of Geography this project would not have been possible. Thank you to Mr Schalk van Heerden and Rhett Calvert for enthusiastic support with the capturing of GPS coordinates.

9. References


