ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) 550–557

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Crowdsourcing geospatial data

Christian Heipke

Institut für Photogrammetrie und GeoInfomation, Leibniz Universität Hannover, Nienburger Street 1, 30167 Hannover, Germany

ARTICLE INFO

Article history:
Received 17 January 2010
Received in revised form
19 April 2010
Accepted 22 June 2010
Available online 12 August 2010

Keywords:
Mapping
Web 2.0
Geo-referencing
User-generated content
Crowdsourcing

ABSTRACT

In this paper we review recent developments of crowdsourcing geospatial data. While traditional mapping is nearly exclusively coordinated and often also carried out by large organisations, crowdsourcing geospatial data refers to generating a map using informal social networks and web 2.0 technology. Key differences are the fact that users lacking formal training in map making create the geospatial data themselves rather than relying on professional services; that potentially very large user groups collaborate voluntarily and often without financial compensation with the result that at a very low monetary cost open datasets become available and that mapping and change detection occur in real time. This situation is similar to that found in the Open Source software environment.

We shortly explain the basic technology needed for crowdsourcing geospatial data, discuss the underlying concepts including quality issues and give some examples for this novel way of generating geospatial data. We also point at applications where alternatives do not exist such as life traffic information systems. Finally we explore the future of crowdsourcing geospatial data and give some concluding remarks.

1. Introduction

For centuries mapping has been accomplished by state organisations. While in the early days, mostly military interests dominated mapping activities, in the 20th century emphasize shifted more and more to civilian National Mapping and Cadastre Agencies (NMCAs), mostly responsible for land administration, infrastructure planning or environmental monitoring. The work was carried out by trained staff based on well defined specifications as well as mapping and quality assurance procedures.

It was only within the last five years or so that a new player has entered the scene: the citizens themselves have started to become involved in mapping their immediate environment. The development was made possible by progress in positioning and navigation, namely portable and low-cost GPS receivers often combined with a camera in a mobile or a smart phone, and in computer science, in particular the emergence of web 2.0 technology (O'Reilly, 2005), combined with broad band communication links. Most of the citizens lack formal training in mapping; furthermore map specifications and quality assurance procedures are often absent in these endeavours. Nevertheless, in many cases the quality of the resulting geospatial data is rather good. It seems thus, that the borders between the once distinct roles of producer, service provider and user of geospatial data are substantially being blurred, as amateurs start doing the job of professional surveyors and distributors, and accomplish it rather well compared to traditional mapping products (see also Sections 2.5 and 3.2).

This development, more apparent in some countries than in others, was partly inspired by the success of projects such as Wikipedia (en.wikipedia.org/wiki/Main_Page), the free online encyclopaedia, and the Open Source software initiative (www.opensource.org). Various terms have been coined for the new way of mapping such as *neogeography* (Turner, 2006; Rana and Joliveau, 2009) and *volunteered geographic information* (Goodchild, 2007). Here we speak of *crowdsourcing*[1] geospatial data, see also Ramm (2008) and Hudson-Smith et al. (2009).

The field is constantly growing and has lots of potential for mapping, mainly due to two reasons: (a) the technology is mature and is there to stay, (b) other means of mapping may be too slow (e.g. for car navigation systems or disaster mapping) or may be too expensive (e.g. updating the United States Geological Survey (USGS) 1:24.000 map scale on a countrywide basis).

In this paper we shortly review the crowdsourcing activities for geospatial data. After explaining the concepts and some of the basic technology, we discuss a few pertinent questions such as (a) which

[1] The term *crowdsourcing* is derived from *outsourcing*, where production is transferred to remote and potentially cheaper locations. Analogously, crowdsourcing describes the concept where potentially large user groups carry out work which is expensive and/or difficult to automate.

E-mail address: heipke@ipi.uni-hannover.de.

0924-2716/$ – see front matter © 2010 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2010.06.005

are the data sources being used for crowdsourcing?, (b) who are the crowds collecting this geospatial data and what motivates them to allocate part of their free time to mapping activities?, and (c) what is the quality of the obtained datasets? We then give a few examples of existing crowdsourcing projects, both from the open and the commercial domain, before concluding with some remarks about future developments.

2. Issues in crowdsourcing geospatial data

2.1. Definitions

As one would expect for a new field, many diverse definitions exist for neogeography, volunteered geographic information and crowdsourcing geospatial data. Following Turner, "Essentially, neogeography is about people using and creating their own maps, on their own terms and by combining elements of an existing toolset. Neogeography is about sharing location information with friends and visitors, helping shape context, and conveying understanding through knowledge of place" (Turner, 2006, p. 3), while Rana and Joliveau note that ".. most researchers would agree on the common keywords that have come to be associated with the works on NeoGeography. These keywords include web 2.0, mashups, public participation, social networking, volunteered geographical information, crowd-sourced data, user-generated content, Open Source maps API [application programming interfaces] and affordable navigational devices, principally GPS data loggers. If one brings together these ideas and tools to represent some geographical information, the outcome falls somewhere in the undefined extent of NeoGeography." (Rana and Joliveau, 2009, p. 2), and Goodchild puts forward that "the widespread engagement of large numbers of private citizens, often with little in the way of formal qualifications, in the creation of geographic information, […] almost always voluntar[il]y […] represent[s] a dramatic innovation that will certainly have profound impacts on […] the discipline of geography and its relationship to the general public. I term this *volunteered geographic information* (VGI), a special case of the more general Web phenomenon of *user-generated content* […]" (Goodchild, 2007, p. 212).

It is thus clear that content generation by large numbers of users, amateurs as far as data collection is concerned, with the help of web technology is a key point of this development. As Goodchild (2008) points out, however, geography (and one may add geospatial sciences as a whole, including photogrammetry, remote sensing and cartography; see e.g. (Heipke, 2004) for a discussion on the respective relations), is far more than data acquisition, as it includes data modelling prior to acquisition, data integration and interpretation with a view to some application.[2] Therefore, we prefer the term *crowdsourcing* to describe data acquisition by large and diverse groups of people, who in many cases are not trained surveyors and who do not have special computer knowledge, using web technology. Subsequently, these data are transferred to and stored in a common computer architecture, e.g. a central or a federated database, or in a cloud computing environment (Leymann, 2009). The subsequent tasks of automatic data integration and processing are very important to generate further information from the acquired data, sometimes also in feedback loops with data acquisition, and provide multiple research challenges. However, these tasks are not central to the discussions of this paper. Rather, the focus is on the data acquisition aspects, and we discuss some forms of automatic processing in chapter 4 only.

2.2. Basic technologies

There are two basic technologies which have made possible the success of crowdsourcing geospatial data. Both are easy to use and have found a widespread distribution in the population (at least in the developed countries, see also Section 4). The technologies are (a) geo-referencing, either using GPS or mobile phone positioning and (b) the web 2.0 development incl. broad band communication.

It is clear that in order to acquire geospatial data a datum must be provided. The easiest way to do so is to use a global datum, and global navigation satellite systems (GNSS) such as GPS and Galileo do exactly that. Another possibility to easily geo-reference acquired data is to digitise them from an existing map or digital orthophotos or aerial image, in which every pixel is at least approximately geo-referenced. The same holds if refinements (e.g. provision of additional attributes) are made for existing geo-referenced map data.

The second decisive development for crowdsourcing was that of the web 2.0 (O'Reilly, 2005). In web 2.0 web sites can be directly linked to databases by APIs. In this way, users can generate and then upload data and thus make it available to the web community, hence the term 'user-generated content'. In a similar way interactive sharing of information, mashing up web sites (i.e. generating new information by embedding a web site, often with geospatial content such as maps, images or videos, into another one) and collaboration on the web have become possible. These developments were based on the principle of interoperability, and thus standards in data formats and services, creating a linked patchwork of a geodata infrastructure (Goodchild, 2007). Based on these possibilities mapping has become an important application of web 2.0. Yet, as of the time of writing (January 2010), to our knowledge there have not been any malicious attacks on crowdsourced geospatial data, although one must fear that such activity will happen as the subject gets more and more headlines.

2.3. Data sources

Two of the data sources used for crowdsourcing were already mentioned: GPS receivers yielding GPS tracks and orthophotos. Besides, other images can be used as well, provided that they can be (or are already) geo-referenced in one way or another. These images can come as aerial or satellite images, images from mobile phones, mobile mapping vans and from UASs (unmanned aircraft systems). The latter three platforms are interesting, because in this way users can acquire their own images of a local neighbourhood and thus do not have to rely on images available on the web. In this case, image orientation should be determined fully automatically to avoid pitfalls. In general, this is already possible given today's state of the art (e.g. (Mayer and Bartelsen, 2008; Agarwal et al., 2009)), in particular since all three platforms can be equipped with a GPS receiver further constraining the solution.

The use of images for crowdsourcing ((Leberl, 2010), uses the term neo-photogrammetry in this domain) triggers another question, namely in how far primary data acquisition (photo capture in this case) and data exploitation (extracting information from these images) should happen simultaneously in order to qualify as a crowdsourcing activity. In the extreme case, it could be argued that everybody uploading his vacation images to a portal such as flickr (www.flickr.com) without any intention to do mapping is still part of a group crowdsourcing geospatial data, and that data mining on the web is a crowdsourcing activity as such.

[2] The same argument also holds for the measurement sciences: obtaining high quality results is not only a question of appropriate hardware (the cost of which may be low in the future), but requires proper design of the whole task, special care in handling the measurement devices, taking additional observations to correct for systematic errors, self-diagnosis to avoid blunders, and a final quality evaluation of the obtained results.

While this view may sound a little far fetched, there is no doubt that crowdsourcing can indeed happen in different, unrelated steps. Along similar lines, mapping may happen without people being aware of it. This is for instance the case when anonymized signals of mobile phones carried by car drivers or pedestrians are being used to map out roads or pedestrian walkways.

Another important data source is local knowledge. It is well known from traditional mapping that operators familiar with the local environment make less errors during data acquisition, have fewer problems in dealing with ambiguities and thus derive higher quality results. Crowdsourcers are often most interested to map their immediate surroundings, which is the area they know best. This inherent benefit of crowdsourcing may be responsible for much of the – perhaps surprisingly good – quality of crowdsourced data (see also below).

Finally, data can be (and have been) donated to crowdsourcing projects, mainly from government authorities; from raw GPS tracks and images to fully labelled vector data.

2.4. One crowd vs. many crowds

A group of people collectively carrying out a mapping project will most probably have something in common besides mapping, they form a socially linked group. Since there are many different social environments, crowdsourcers must be expected to form different groups with different needs and thus different ways to keep them motivated. For instance, while some groups would certainly be more active in reporting omissions and errors in datasets, others will be more likely to do a complete mapping project.

We continue with describing different mapping crowds. In this discussion we draw on material which was elaborated during a recent EuroSDR workshop (Streilein et al., 2010), keeping in mind that the topic of the workshop was crowdsourcing as a support for NMCAs, not as a separate endeavour. Also, we mention that the crowdsourcers can form open groups and would then be regarded as partners in the mapping project, or they form closed groups which work as employees for payment, e.g. according to the scheme of Amazon's Mechanical Turk (www.mturk.com/mturk/welcome).

The groups identified during the EuroSDR workshop are:

- *Map lovers*, a small group who produce trustable and very valuable data. An example may be a retired military officer who has relied on correct maps throughout his professional life. If such a person finds an error in the map, he may be very willing to let authorities know.
- *Casual mappers*, e.g. hikers, bikers, mountaineers etc. Casual mappers partly overlap with the map lovers but can be distinguished by the fact that they are only willing to spend a relatively low effort for mapping. Apart, they would probably rather upload new data than report errors.
- *Experts* are active people and leading map users in organisations like mountain rescue, fire brigades, civil protection, traffic guides, etc. They are motivated by the feeling that they may make their own life easier, they can contribute with very valuable and trustworthy data, both old and new.
- *Media mappers* are potentially large groups, activated sporadically by regional up to international media campaigns. These are mostly once-off mappers especially motivated by competitions, mapping parties, etc. The contributions are limited in time and extent and a big initial effort for the campaign is needed.
- *Passive mappers*: Mobile phones with or without GPS positioning are a prime example of the new technologies that enable passive, and sometimes unaware, data collection about position, time, direction and speed of individuals. While personalised use of the data is prohibited in most cases, anonymized collection of this data is legally possible. The resulting data needs to be further analysed and interpreted for many mapping tasks. For example, car navigation needs an attributed graph data structure in order to derive routing information (see Fig. 1).
- *Open mappers*: Besides the passive mappers the open mappers are by far the largest group, and their number is constantly growing. Therefore, they form an increasingly important group, spending a significant amount of time and effort to build open datasets. They are motivated by contributing (and using) good public data, which currently hampers their cooperation with most NMCAs. The enhancement and simplification of the associated Open Source mapping tools developed e.g. by the Open Source Geospatial Foundation (www.osgeo.org), will help to further grow this group.
- *Mechanical turks*: The Amazon Mechanical Turk (see above) is a crowdsourcing marketplace and service where humans can contribute to posted tasks for a monetary payment. Clearly, the motivation of this group is money, which sets them aside from the other mentioned groups. Validation of geospatial data could for instance be organised using a mechanical turk service by superimposing information from a geospatial database with that from an aerial or satellite image.[3]

Discussing the issue of motivating people in a group to start or continue working in a web-based collaborative environment leads into interesting relations to social networks on the web such as Facebook or MySpace, but is far beyond the scope of this paper. Instead, we restrict ourselves to a few comments only. According to the 90:9:1 rule observed by Nielsen (2006) for open contribution systems, 90% of the users only consume the information, 9% contribute occasionally, and only 1% of the user community is constantly active with contributing information. Nielsen argues that this observation is valid for many examples, and sometimes with even more unbalanced ratios. He also suggests a few measures to increase active participation:

- make it easy to contribute, without technical, logistic, legal or intellectual barriers,
- make participation a side effect (this is the principle of passive mapping),
- edit, don't create (provide templates for jobs to be done),
- reward, but don't over-reward active participants, otherwise they may dominate the whole project even more,
- promote quality contributors, e.g. by reputation ranking (see e.g. Sayda, 2005). This possibility is well known from shopping activities on the web, where it is designed to build trust.

It may be added that making participation popular (so one is more inclined to talk about it to his friends) and making it to be fun, e.g. in a competitive game (e.g. Matyas, 2007) or a geocaching activity, will certainly also help as will local pride ("I want my own house/neighbourhood/way to work to be present in the dataset") and a possibility of self-promotion.

2.5. Quality of crowdsourced geospatial data

It is sometimes argued that a major goal of crowdsourcing is the social experience to share a common activity (here mapping) and to have fun doing so. Nevertheless, many mapping applications (and thus also quality requirements) are shared between users of crowdsourced and those of traditionally acquired datasets. Since

[3] To cite another example (although an unsuccessful one, (Havercroft, 2010)), the remains of US explorer Steve Fossett were searched for on high resolution satellite images shortly after his plane had crashed in September 2007 using a beta version of Amazon's Mechanical Turk service. Within about one month 50.000 people took part in the exercise.

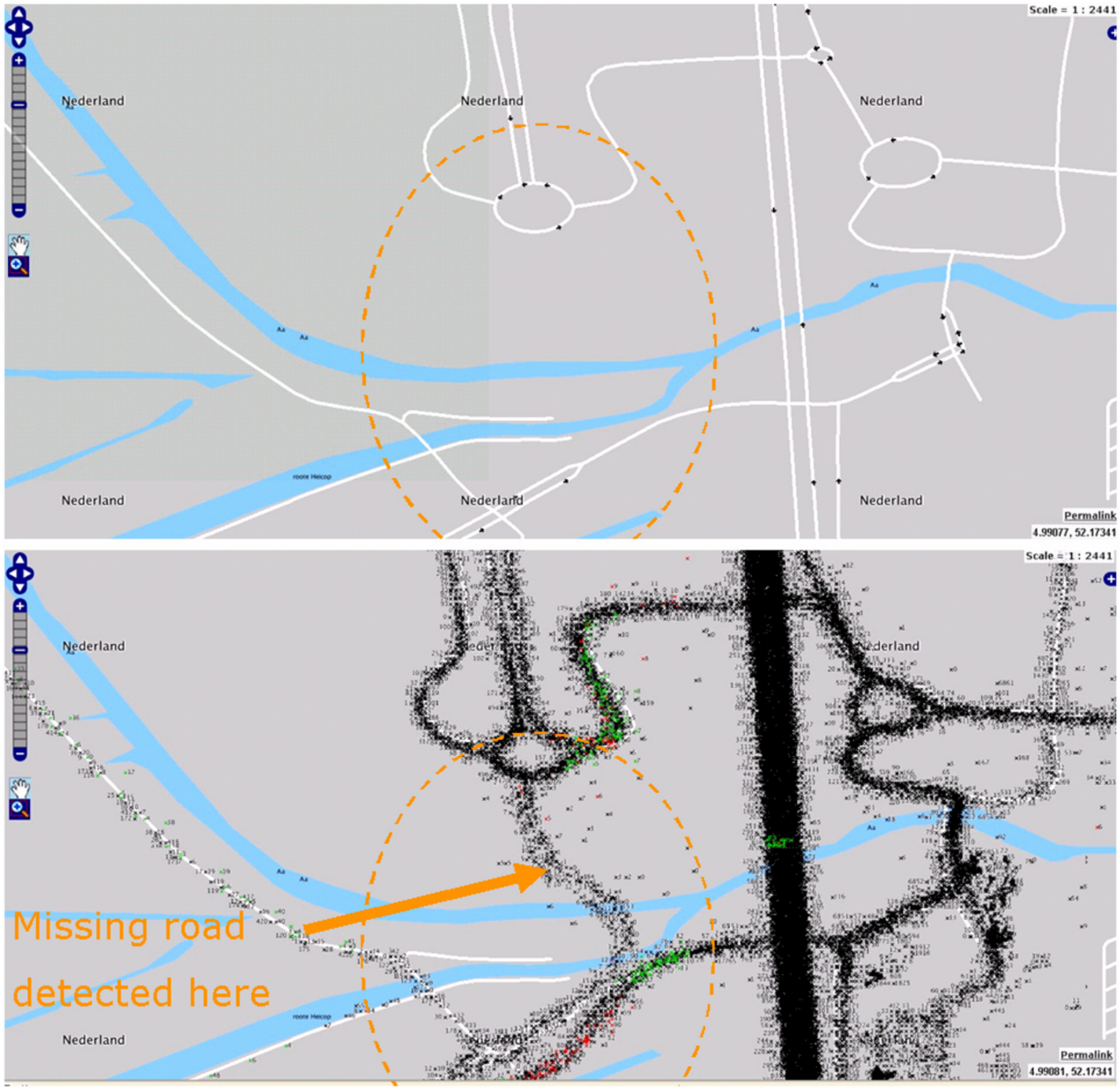

Fig. 1. Example of a result from passive mapping. Upper: existing map data, lower: superimposition with data from GPS tracks derived from cars. The missing road in the centre is clearly visible, however the individual tracks need to be aggregated into an attributed graph structure for use in routing applications (© Tele Atlas).

information collection in crowdsourcing is carried out largely by untrained users, which is in clear contrast to mapping activities led or coordinated by NMCAs, quality issues have been a primary point of debate since crowdsourcing results started to appear. While quality is sometimes described as "fitness for use", this definition is application dependent. We therefore approach the topic from a description of the quality parameters inherently linked to the datasets.

In this context, a few observations shall be made here. The topic is picked up again in Section 3.2, when the quality of the OpenStreetMap dataset is discussed:

- in traditional mapping as well as in crowdsourcing the same set of standardised quality parameters (completeness, up-to-dateness, relative and absolute geometric accuracy, attribute correctness, topological correctness, logical consistency) must be considered, see ISO/TC211 (www.isotc211.org) and Joos (2000);
- NMCAs (and often also successful commercial companies) are considered as trustworthy organisations by many people based on past experience and since in most cases they have well documented mapping specifications as well as workflow and quality assurance procedures. Therefore, authoritative geospatial data and those provided by car navigation companies (Tele Atlas, Navteq) and providers of virtual globes (e.g. Google, Microsoft) have an inherent advantage as far as trust by the general public goes; this trust relates to all quality parameters. On the other hand, crowdsourced data comes with the advantage of local knowledge (see above), the results are sometimes labelled as *accredited*. Furthermore, since crowdsourcing is carried out by many people, repetitions will occur much more frequently, and from a statistical point of view one can expect to have a rather low rate and size or errors[4];
- most, if not all, NMCAs have a mandate to cover the whole territory they are responsible for in an equal manner. Therefore,

[4] Raymond (2002) summarises this fact in what he calls Linus's law: "Given enough eyeballs, all bugs are shallow". The name of the law is derived from that of Linus Torwalds, the initiator of the Linux Kernel.

quality parameters tend to be homogeneous across space, time and the thematic domain, although differences may occur between urban and rural areas. The distribution of quality parameters, and in particular spatial, temporal and thematic completeness, is much more questionable for crowdsourced geospatial data (see also Brassel et al., 1995; Maué and Schade, 2008);

- while GPS has indeed revolutionized surveying and mapping and has brought possibilities for very accurate measurements to the lay person, the low-cost measurement devices normally used in crowdsourcing limit at least the currently attainable geometric accuracy. For example, deformation surveys of tall bridges or dams requiring millimetre accuracy over large distances is certainly out of the range of crowdsourcing, as is the determination of gradual height changes due to subsidence in mining areas.

3. Examples of crowdsourcing systems

3.1. Overview

A large number of projects for crowdsourcing geospatial data have been established over the last few years. Some of the more prominent examples include OpenStreetMap and Wikimapia, which follow the ideas of the Open Source initiative and produce openly available data, as well as the commercial systems Google Map Maker and TomTom's HD Traffic™ real-time traffic information.

A system in the domain of disaster mapping, already going back more than 10 years, is a portal for earthquake mapping called "Did you feel it?" established by the USGS Earthquake Hazards Program (earthquake.usgs.gov/earthquakes/dyfi, Wald et al., 1999) which collects experiences from people having been affected by an earthquake. Poser et al. (2009) investigate possibilities to asses the quality of data acquired from the affected population in a case study on flooding.

Also a number of NMCAs such as IGN-France (Viglino, 2010), the Ordnance Survey (Havercroft, 2010) and swisstopo (Guélat, 2010) have started to experiment with establishing portals for crowdsourcing. In addition, a number of groups have set up smaller projects for special purposes, e.g. for counting birds over Christmas (www.audubon.org/Bird/cbc), the roots of which go back more than 100 years! The last example mentioned in this overview is Geo-Wiki (Fritz et al., 2009; www.geo-wiki.org), a project to validate global land cover data based on satellite images available on Google Earth.

3.2. OpenStreetMap

Probably the most prominent crowdsourcing project for geospatial data based on the ideas of the Open Source initiative is OpenStreetMap (www.openstreetmap.org). The goal of this project, founded in 2004 by Steve Coast, then with University College London, is to enable free access to current digital geographical information across the world. At the time of writing (January 2010) statistics in the OpenStreetMap website report 200.000 mappers, up from 50.000 in August 2008 and 100.000 in March 2009. The number of participants is thus increasing rapidly and at an accelerating speed. It is estimated that approximately 10% of the participants are active on a monthly basis (Ramm, 2008).

Mapping is instantaneous and is accomplished as soon as a crowdmapper finishes his job. Data acquisition is based on the sources mentioned above. OpenStreetMap has been given access to aerial images from Yahoo! and was also able to import the TIGER dataset from the US Census Bureau. Mapping is guided by a set of predefined features and attributes, but as crowdmappers map everything they think is of interest, a common feature catalogue does not exist, and the degree of geometric and attribute detail may vary considerably from one area to another. As additional features are being included into the dataset, existing viewers and other tools may have to be adapted to exploit the new features. A number of services have been created around OpenStreetMap, an example is OpenRouteService (www.openrouteservice.org, Schmitz et al., 2009). Of particular interest for data quality is OpenStreetBugs (openstreetbugs.appspot.com), a tool designed for the crowdmappers to mark and eventually correct detected errors in the dataset. Currently, OpenStreetMap is a two-dimensional dataset, however first attempts to import a Digital Terrain Model have already been reported (Zipf, 2009; personal communication).

A particular example of the efficiency of the OpenStreetMap approach was the heavy earthquake in Haiti in January 2010. Within days of the disaster the OpenStreetMap community initiated a mapping project based on current satellite images to provide geospatial data to rescue workers and the island inhabitants at large. Updating took place on an hourly basis, and a detailed city map of the capital Port-au-Prince containing bridges, functioning infrastructure and damaged buildings was finished after only a few days (see Fig. 2).

The quality of the OpenStreetMap data has been investigated in a number of studies. Hacklay (in press) reported about 80% coverage and a geometric accuracy of about 6 m for major roads in the London area compared to Ordnance Survey Meridian 2 data based on the buffer method (Goodchild and Hunter, 1997; Heipke et al., 1998) and a buffer width of 20 m. On a national comparison, nearly 30% of England had been mapped by volunteers in the four years since 2004 with a distinct difference between cities (detailed datasets, in particular for London and Birmingham) and the countryside (much less activity). Since March 2008 completeness has significantly gone up and stands at about 65% without considering attribute information as of October 2009 ((Hacklay, 2009a), see also Fig. 3). As is to be expected, completeness still depends on population density, and it also depends on the level of income: more affluent areas are better mapped then more deprived areas (Hacklay, 2009b). Similar investigations confirm these numbers for Athens (Kounadi, 2009) and for areas in Germany (Zipf, 2009; personal communication), other countries seem to be less mapped when taking a look at the data available on the web.

3.3. *HD* Traffic*™*

The car navigation industry is a very prominent example for a user who needs instant information, because current information considerably increases the value for drivers. The system HD Traffic™ from TomTom (www.tomtom.com/hdtraffic) works with probe data and is based on anonymized location, direction of movement and speed information from active mobile phone calls available from mobile phone companies and on GPS position information from the installed TomTom user base currently linked to the service; the processing chain is patented (Atkinson et al., 2004). Phone data are first filtered to separate data of car drivers from those of other users. The system then aggregates the filtered results and combines them with road map data and speed profiles for individual roads collected in the past to produce live traffic and navigation information. Information is updated every three minutes.

HD Traffic™ can be regarded as a crowdsourcing system, where the crowd is part of the group of passive mappers. The quality of the service is obviously linked to the number of people with mobile phones and TomTom users travelling on the roads. Thus, information on secondary roads may not be as useful as that on main roads and motorways. At the time of writing HD Traffic™ is available in a number of countries in Europe such as the Belgium, France, Germany, the Netherlands, Portugal, Switzerland and the UK.

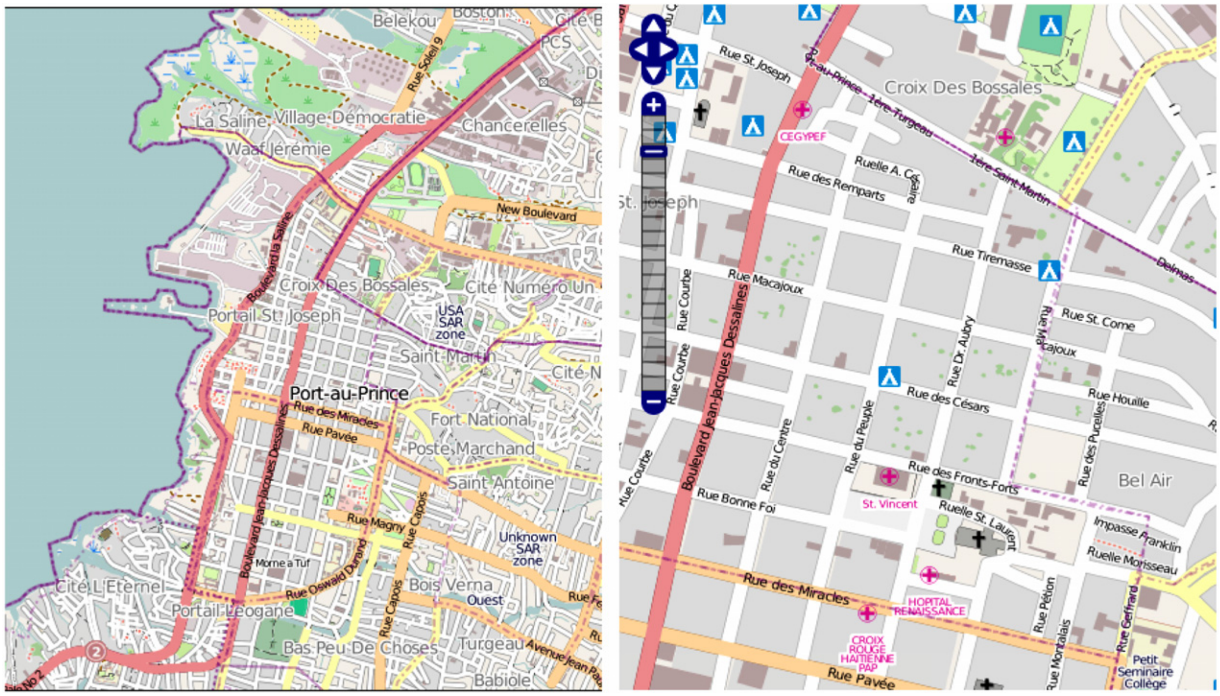

Fig. 2. Two views of the OpenStreetMap city map of Port-au-Prince, Haiti, generated after the Jan-2010 earthquake; left: overview, right: detail (http://www.openstreetmap.org/?lat=18.5531&lon=-72.3063&zoom=13&layers=B000FTF).

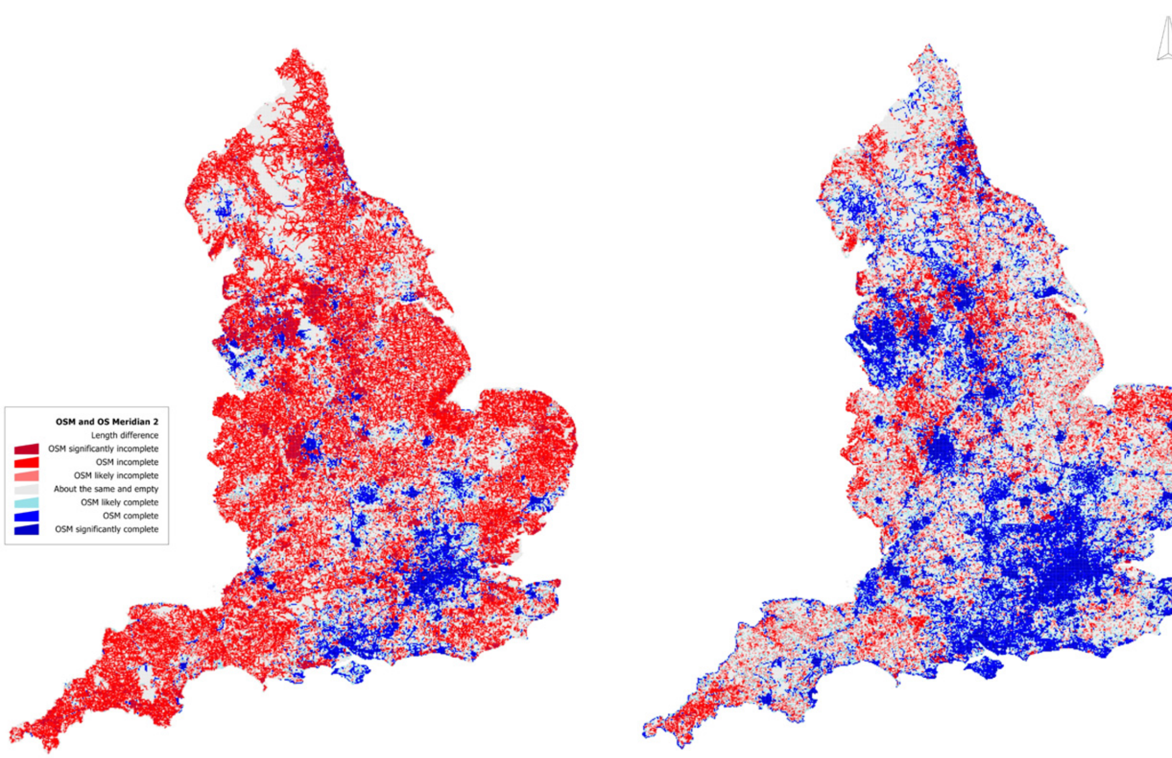

Fig. 3. Completeness of OpenStreetMap data compared to Ordnance Survey Meridian 2 data. Left: March 2008, right: October 2009. Red signifies areas where OpenStreetMap data are significantly incomplete, in blue areas OpenStreetMap data are significantly complete as judged by a comparison of length.
Source: From Hacklay (2009a).

4. Further developments and conclusions

In the short term geospatial crowdsourcing will certainly see improvements in the workflow and the human-computer-interface of crowdsourcing systems making it easier for the individuals to contribute information. Also, data models will become more sophisticated, perhaps including different levels for different kinds of processing (as e.g. suggested for car navigation by Brenner, 2008) and the third dimension.

Automatic processing of spatial data involving computer vision (e.g. Forsyth and Ponce, 2003), pattern recognition and machine learning (e.g. Bishop, 2006), which has already been a hot topic of research for traditionally acquired geospatial data, will obviously also be applicable and will be very useful when dealing with crowdsourced data, as already mentioned in Section 2.1. Data integration algorithms on the geometric as well as the semantic level (e.g. Gösseln and Sester, 2005; Matyas, 2007) will be refined to improve the results; the geometric aggregation of individual GPS tracks from passive mapping to generate attributed graph data structures as thus routable road information can be largely carried out in an automatic fashion already today (Cao and Krumm, 2009). Digitisation from images will increasingly be supported by automatic or semi-automatic image analysis tools, e.g. for assessing available vector data (e.g. Gerke and Heipke, 2008) or for extracting individual objects (e.g. Wolf(né Straub) and Heipke, 2007).

In particular for the open datasets a number of non-technical questions arises: how can the projects be made sustainable enough, that they can continue to exist once the initial hype has passed? Can they generate the required support, for instance for the necessary computer infrastructure, for the maintenance of the data bases, but also for access to the GNSS infrastructure? Is there a way for cooperation between the open mappers and NMCAs for mutual benefit, despite the clearly different business models, licensing constraints, questions around quality standards and accountability (for the latter issue: could a temporary additional attribute called "crowdsourced", to be substituted at a later stage after authoritative verification, be of some help)? What are financial implications for NMCAs and for commercial mapping companies, if more and more mapping applications are based on open data in the future? Is the provision of services on top of these data a profitable new business model for them? Addressing these questions at this point in time involves a substantial amount of speculation, but the answers are needed and will certainly be given sooner or later.

In the longer term other issues become more important: Technology will develop further and will offer new possibilities to cooperate across the web. In what they call Web^2 or Web Square, O'Reilly and Batelle describe what the web may look like in five years from now (O'Reilly and Batelle, 2009, note how often the authors refer to location and positioning in their predictions): People, as well as other real-world objects, cast information shadows in cyberspace (see also Kuniavsky, 2009), and thus related information can be collected by automatically processing this data available on the web (see discussion on passive mappers above). In this regard, as mobile phones are augmented with additional sensors and start communicating and cooperating with each other, geosensor networks, that is networks of connected and potentially mobile sensors often directly linked to the web (Stefanidis and Nittel, 2004), will play an increasingly dominant role. These sensors report position and time together with additionally collected information, and humans can obviously not only serve as platforms for these sensors, but also as sensors themselves (see e.g. Goodchild, 2007; Heipke and Sester, 2008).

When combining the large amounts of raw sensor data thus collected with the growing potential of image analysis, pattern recognition and machine learning, future mapping may only involve taking images of objects of interest; turning the image into meaningful information is then accomplished employing real-time automatic image interpretation, perhaps supported by the photographer providing a written annotation or giving a few oral hints and a handwriting and speech recognition software translating these annotations and hints into computer readable information. Visualisation can then also take place in real time by superimposing the results to the image adopting appropriate augmented reality approaches, perhaps together with additional information from a connected database. While this is certainly a long-term vision, with many hurdles unknown as yet, it seems that with the development of the web to date we have taken this direction, and reflecting on the progress of mapping reviewed in this paper we have already accomplished a considerable part of the way.

A final remark, already observed by Goodchild (2007) and others, shall be made: crowdsourcing using internet technology obviously requires internet technology. While globalisation has decreased distances between different parts of the world in many metrics, access to the internet is not necessarily one of them, a phenomenon sometimes called the digital divide. Currently, a fundamental limitation for mapping the whole world via crowdsourcing approaches is the lack of affordable technology in many parts of the world. Another issue with potentially large implications is the fact that most web services, where available, can only be used with sufficient knowledge of the English language and the Latin alphabet. These are constraints which need to be seriously taken into account when approaching and trying to solve global issues based on globally collected geospatial data with global support.

Acknowledgements

Many thanks are due to André Streilein and his team for providing me a copy of the report on the EuroSDR Workshop *Crowdsourcing for updating national databases*, held in Berne, Switzerland in August 2009. I am also grateful to Monika Sester for her valuable comments on an early draft of the manuscript which considerably improved the paper.

References

Agarwal, S., Snavely, N., Simon, I., Seitz, S.M., Szeliski, R., 2009. Building Rome in a day, Proceedings International Conference on Computer Vision, ICCV, Kyoto, Japan, (accessed 02.01.10).

Atkinson, I., Adam, B., Dixon, J., 2004. Traffic Monitoring System, European Patent EP 1 348 208 B1, www.freepatentsonline.com/EP1348208.pdf (accessed 02.01.10).

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, Berlin, p. 739.

Brassel, K., Bucher, F., Stephan, E., Vckovski, A., 1995. Completeness. In: Guptill, S.C., Morrison, J.L. (Eds.), Elements of Spatial Data Quality. Elsevier, pp. 81–108.

Brenner, C., 2008. Geosensor networks—the case of navigation and driver assistance systems. In: Heipke C., Sester M. (Eds.), 2008: Geosensor Networks, Proceedings EuroSDR & ISPRS Workshop Hannover, EuroSDR Official Publication No. 54, bono.hostireland.com/~eurosdr/km_pub/no54/html/Geosensor_WS/index.htm (accessed 02.01.10).

Cao, L., Krumm, J., 2009. From GPS traces to a routable road map. In: Proceedings 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM GIS '09. ACM Publisher, Seattle, ISBN: 978-1-60558-649-6, pp. 3–12.

Forsyth, D.A., Ponce, J., 2003. Computer Vision, A Modern Approach. Prentice Hall, Upper Saddle River, p. 693.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., Obersteiner, M., 2009. Geo-Wiki.Org: the use of crowdsourcing to improve global land cover. Remote Sensing 1 (3), 345–354.

Gerke, M., Heipke, C., 2008. Image based quality assessment of road databases. International Journal of Geoinformation Science 22 (8), 871–894.

Goodchild, M., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal 69 (4), 211–221.

Goodchild, M., 2008. Assertion and authority: the science of user-generated geographic content, In: Navratil G. (Ed.), Proceedings of the Colloquium for Andrew U. Frank's 60th Birthday. GeoInfo 39. Department of Geoinformation and Cartography, Vienna University of Technology, 18 p.

Goodchild, M., Hunter, G., 1997. A simple positional accuracy measure for linear features. International Journal of Geographical Information Science 11 (3), 299–306.

Gösseln, G.V., Sester, M., 2005. Change detection and integration of topographic updates from ATKIS to geoscientific datasets. In: Agouris, P., Croitoru, A. (Eds.), Next Generation Geospatial Information. In: ISPRS Book Series, Taylor & Francis Group, London, pp. 69–80.

Guélat, J.-C., 2010. Integration of user generated content into national databases—revision workflow at swisstopo, In: Streilein A., Schaper J., Kellenberger T., Koenig T., Guélat J.-C., 2010: crowdsourcing for updating national databases, Proceedings EuroSDR Workshop Berne, EuroSDR Official Publication No. 57, http://www.eurosdr.net/workshops/crowdsourcing_2009 (accessed 02.01.10).

Hacklay, M., 2008. How good is volunteered geographical information? a comparative study of OpenStreetMap and Ordnace Survey datasets. In: Environment and Planning B: Planning and Design (in press) Pre-print available at: www.ucl.ac.uk/~ucfamha/OSM%20data%20analysis%20070808_web.pdf (accessed 02.01.10).
Hacklay, M., 2009a. OpenStreetMap and Ordnance Survey Meridian 2— progress maps. PoVeSham personal blog Nov-14-2009, povesham.wordpress.com (accessed 02.01.10).
Hacklay, M., 2009b. The digital divide of OpenStreetMap, PoVeSham personal blog Dec-28-2009, povesham.wordpress.com (accessed 02.01.10).
Havercroft, M., 2010. Crowdsourcing—some experiences and thoughts, In: Streilein A., Schaper J., Kellenberger T., Koenig T., Guélat J.-C., 2010: crowdsourcing for updating national databases, Proceedings EuroSDR Workshop Berne, EuroSDR Official Publication No. 57, http://www.eurosdr.net/workshops/crowdsourcing_2009 (accessed 02.01.10).
Heipke, C., 2004. Some requirements for geographic information systems: a photogrammetric point of view. Photogrammetric Engineering & Remote Sensing 70 (2), 185–195.
Heipke, C., Mayer, H., Wiedemann, C., Jamet, O., 1998. External evaluation of automatically extracted road axes. Photogrammetrie – Fernerkundung - Geoinformation 2, 81–94.
Heipke, C., Sester, M., 2008. Geosensor Networks, Proceedings EuroSDR & ISPRS Workshop Hannover, EuroSDR Official Publication No. 54, bono.hostireland.com/~eurosdr/km_pub/no54/html/Geosensor_WS/index.htm (accessed 02.01.10).
Hudson-Smith, A., Batty, M., Crooks, A., Milton, R., 2009. Mapping for the masses: accessing web 2.0 through crowdsourcing. Social Science Computer Review 27 (4), 524–538.
Joos, G., 2000. Zur Qualität von objektstrukturierten Geodaten. In: Schriftenreihe Studiengang Geodäsie und Geoinformation, Universität der Bundeswehr München, Heft 66, Neubiberg, 2000.
Kounadi, O., 2009. Assessing the quality of OpenStreetMap data. M.Sc. Thesis, Dep. of Civil, Environmental and Geomatics Engineering, University College London, 81 p. homepages.ge.ucl.ac.uk/~mhaklay/pdf/Rania_OSM_dissertation.pdf (accessed 02.01.10).
Kuniavsky, M., 2009. Smart things—an outline, public notebook Feb-8-2009, www.orangecone.com/archives/2009/02/smart_things_an.html (accessed 02.01.10).
Leberl, F., 2010. Time for neo-photogrammetry? GIS Development 14 (2), 22–24.
Leymann, F., 2009. Cloud computing: the next revolution in IT. In: Fritsch, D. (Ed.), Photogrammetric Week'09. Wichmann Publisher, pp. 3–12.
Matyas, S., 2007. Collaborative spatial data acquisition—a spatial and semantic data aggregation approach. In: Proceedings, 10th AGILE International Conference on Geographic Information Science, Aalborg, 6 p. (on CD-ROM).
Maué, P., Schade, S., 2008. Quality of Geographic Information Patchworks. In: Bernard L., Friis-Christensen A., Pundt P., Compte I. (eds.), 11th AGILE International Conference on Geographic Information Science, Girona, 8 p. (on D-ROM).
Mayer, H., Bartelsen, J., 2008. Automated 3D reconstruction of urban areas from networks of wide-baseline image sequences, the international archives of the photogrammetry. Remote Sensing and Spatial Information Sciences 37 (B5), 633–638.
Nielsen, J., 2006. Participation inequality: encouraging more users to contribute, Alertbox personal blog Oct-9-2006, www.useit.com/alertbox/participation_inequality.html (accessed 02.01.10).
O'Reilly, T., 2005. What is web 2.0: Design patterns and business models for the next generation of software, O' Reilly blog, oreilly.com/web2/archive/what-is-web-20.html (accessed 02.01.10).
O'Reilly, T., Batelle, J., 2009. Web squared: web 2.0 five years on, O' Reilly blog, assets.en.oreilly.com/1/event/28/web2009_websquared-whitepaper.pdf (accessed 02.01.10).
Poser, K., Kreibich, H., Dransch, D., 2009. Assessing volunteered geographic information for rapid flood damage estimation. In: Haunert J.-H., Kieler B., Milde J. (Eds.), 12th AGILE International Conference on Geographic Information Science, Hannover, 9 p. (on CD-ROM).
Ramm, F., 2008. Crows sourcing geodata, In: Proceedings FOSSGIS 2008, Freiburg, www.geofabrik.de/media/2009-09-08-crowdsourcing-geodata.pdf (accessed 02.01.10).
Rana, S., Joliveau, T., 2009. NeoGeography: an extension of mainstream geography for everyone made by everyone?, Editorial. Journal of Location Based Services 3 (2), 75–81.
Raymond, E.S., 2002. The cathedral and the bazaar, version 3.0, (accessed 02.01.10).
Sayda, F., 2005. Involving LBS users in data acquisition and update. In: Toppen F., Painho M. (Eds.), 8th AGILE International Conference on Geographic Information Science, Estoril, 10 p. (on CD-ROM).
Schmitz, S., Zipf, A., Neis, P., 2009. New applications based on collaborative geodata—the case of routing. Indian Cartographer 28, 1–6.
Stefanidis, A., Nittel, S. (Eds.), 2004. GeoSensor Networks. CRC Press, p. 296.
Streilein, A., Schaper, J., Kellenberger, T., Koenig, T., Guélat, J.-C., 2010. Crowdsourcing for updating national databases. In: Proceedings EuroSDR Workshop Berne, EuroSDR Official Publication No. 57, http://www.eurosdr.net/workshops/crowdsourcing_2009 (accessed 02.01.10).
Turner, A., 2006. Introduction to Neogeography. O'Reilly Media Publisher, p. 54.
Viglino, J.-M., 2010. Handling partner's feedbacks through the web, In: Streilein A., Schaper J., Kellenberger T., Koenig T., Guélat J.-C., 2010: crowdsourcing for updating national databases. Proceedings EuroSDR Workshop Berne, EuroSDR Official Publication No. 57, http://www.eurosdr.net/workshops/crowdsourcing_2009 (accessed 02.01.10).
Wald, D.J., Quitoriano, V., Dengler, L.A., Dewey, J.W., 1999. Utilization of the internet for rapid community intensity maps. Seismological Research Letters 70 (6), 680–697.
Wolf(né Straub), B.M., Heipke, C., 2007. Automatic extraction and delineation of single trees from remote sensing data. Machine Vision and Applications 18 (5), 317–330.