An outlook for OpenStreetMap in GIScience: experiences, research and applications

Peter Mooney

This volume has presented "*OpenStreetMap in GIScience: experiences, research and applications*" with a collection of experiences and research carried out with OpenStreetMap as the central and core theme. The volume has sought to build a firm foundation to highlight research work focused on OpenStreetMap. This was one of our original goals when we set out at the beginning of the editorial process. This is, to the best of our current knowledge, the first academically produced volume of it's kind which focuses exclusively on OpenStreetMap. Approximately one decade on from the birth of OpenStreetMap in 2004 this volume appears at the most opportune of times. OpenStreetMap has emerged from one of the most tumultuous decades in ICT and possibly in the history of human communication. In the decade where ICT social media, ubiquitous computing and the Internet of Things emerged OpenStreetMap arguably now proudly stands as one of the best examples of crowd and volunteered-based innovation of this time. It's past has been remarkable and the future for OpenStreetMap is bright.

The chapters in the volume have responded to various issues and criticisms aimed at crowdsourced mapping by demonstrating OpenStreetMap's current value and future potential. The cartographic data generation and map making industry has been fundamentally altered by the ICT advancements of the last decade or so. The mapping experience is being transformed (Dodge and Kitchin, 2013). This foundation can be used to support a platform from which future research can be launched. These are exciting times in GIScience. Our edited volume demonstrates that OpenStreetMap is now firmly established on the GIScience research agenda. At the time of writing OpenStreetMap data is increasingly being used as a source of geospatial data for researchers in a wide range of topic areas. These include diverse applications such as the use of OpenStreetMap data for hydraulic modelling in data scarce floodplain areas (Schellekens et al, 2014), to estimating associations between proximity to green spaces and surrounding greenness and pregnancy outcomes (Agay-Shay et al, 2014), to developing models for increased efficiency in road-network warning message dissemination (Fogue et al, 2013)

Many of the key research questions in VGI as outlined by authors such as Mooney et al (2013) are prevalent and relevant to OpenStreetMap. The purpose of this short outlook chapter is to present a picture of where we, as the academic and research community, go from here with OpenStreetMap research. Our edited volume demonstrates that GIScience research on OpenStreetMap has matured. However we urge caution. It is important that the potential generated by this research is maintained and developed going forward. To remain consistent in how we have presented the chapters we will discuss the outlook for OpenStreetMap under the same headings as we have used in the volume.

Data Management and Quality

Amongst the greatest concerns in using VGI to geographers and GIScientsts emerge from issues related to quality assurance of VGI. Is VGI of sufficient quality that it is fit for a specific purpose or use? Many geographers and GIScientists today begun their professional academic training amongst the practices of centralised, scientific, commercial and industrialised geographic information production. These processes were carried out by recognised professionals. Quality assurance procedures were implemented through all of the organisational or professional structures where the geographic information and data were produced. VGI and projects such as OpenStreetMap generate geographic information and data outside these type of organisational and professional structures. Because projects like OpenStreetMap generate geographic information and data in a different setting to established professional structures there are challenges in building quality assurance mechanisms capable of addressing the perceived weaknesses and lack of quality and accuracy in VGI.

There is an abundance of literature where academics and researchers have tackled the quality assurance problem in OpenStreetMap. Studies have shown that in comparison with professionally collected geographical data, such as that from National Mapping Agencies, OpenStreetMap is of equivalent accuracy and in some cases displays better geographical, temporal and attribute accuracy. One of the limitations of these studies surrounds the heterogeneity of OpenStreetMap. Urban areas and agglomerations are usually subject to greater mapping efforts and coverage in OpenStreetMap than rural areas. Some authors have demonstrated heterogeneity in OpenStreetMap related to the socio-economic characteristics of regions within larger areas of population.

However, issues remain. Authors such as Dodge and Kitchin (2013) see VGI projects like OpenStreetMap as perpetually unfinished mapping projects. There are always latent concerns amongst academics and professional users that “something untoward will happen” and data within OpenStreetMap will be damaged, deleted or updated incorrectly immediately rendering it unfit for most application purposes. The greatest challenge faced by OpenStreetMap as it moves forward into it’s second decade is conquering negative perceptions built up from the influence of decades of the established professional production of geographic information and recent commercial interests generating geographic information and services.

The academic and research community must continue to investigate methods for demonstrating the quality of OpenStreetMap data and information and assessing its fitness for purpose for various applications. This is of course a very extensive research mandate. Research efforts will need to be channelled into priority areas such as:

- semantics and interpretation of contributor supplied metadata on primitive OpenStreetMap objects

- geographic accuracy of OpenStreetMap data originating from a heterogenous set of input sources (smartphones, bulk import, on-screen digitizing, etc)
- update and refresh rates of attribute metadata and vector data
- conflation of OpenStreetMap data with other VGI datasets

The Social Context

Many studies have begun to appear which try to understand why, how, where and when citizens contribute to OpenStreetMap. This is not limited to contribution in the form of map edits but contribution in a wider form: software development, website and wiki development, etc. Contributors to OpenStreetMap share what Lin (2014) calls a rather strong shared identity of being an 'OpenStreetMapper'. A sense of community has played an important role in motivating these contributors. OSM can empower individuals and communities. As we have mentioned in this chapter, and have seen in the chapters contained in this volume, there are continuous efforts to understand the emerging practices and methodologies of crowdsourced mapping and geodata collection. This leads us to ask social and critical geographers to consider how we can understand and reveal the social construction of OSM over time.

What are the social mechanisms which have connected thousands of people who have the appropriate interests and resources so that they can collectively contribute to OpenStreetMap? Collective action (Poteete and Ostrom, 2004), as witnessed in OpenStreetMap, has arisen where groups of especially interested, motivated and resourceful individuals are in some way socially connected to each other. However little is really known about the effects of the social ties within the OpenStreetMap community (Mooney and Corcoran, 2013b). Even less is known about the thousands of contributors to the project who *only stay for a while* and leave after contributing maybe only a small handful of edits. Sustaining contributor involvement is also crucial for the long-term future of OpenStreetMap. What are the best ways OpenStreetMap can recruit new contributors whilst ensuring the existing ‘OpenStreetMappers’ remain actively involved and engaged in the project? Gender divisions which are appearing on the GeoWeb and within VGI must be addressed as a matter of urgency (Stephens, 2013).

Network Modelling and Routing

With the smartphone or smartdevice in our pocket people are using location-based services in ever increasing numbers throughout their everyday life. The concepts of directions, navigation and routing are becoming embedded in our social networking applications, web-searches and general smartdevice usage. We have seen examples in this volume where OpenStreetMap is being used as the data source for the development of advanced network models and route finding algorithms. References within these chapters further demonstrate the influence the availability of OpenStreetMap data is having on the development of network models and routing algorithms for Location-based Services and other Internet-based and smart device-based applications. There is still some way to go in regards to addressing and location-based identification in OpenStreetMap. There are millions of building objects in the OpenStreetMap database without any usable addressing information attributed to them. This is something that needs to be considered by the OpenStreetMap community. But perhaps there are opportunities for the academic and research community to use data conflation techniques to extract accurate addresses for these building objects from other openly available geographic datasets or from social media such as Twitter, Foursquare and Flickr (Quercia and Saez, 2014). Addressing and improved location-based identification must be built into OpenStreetMap. The experience and expertise of the academic community from Computer Science, Networking, Mathematics and GeoComputation can assist in driving OpenStreetMap forward in this domain.

This type of collaboration and coordination provides an opportunity for better coordination or cooperation between traditional GIScience processes and VGI. Areas of mutual benefit can be found such as networking modelling and routing. At this stage in our technological life in GIScience it is no longer satisfactory to consider VGI part of a growing cohort of ‘disruptive technologies’. As work presented in this volume has demonstrated OpenStreetMap and other VGI data sources are being considered as an integrated part of the GIScience technological landscape.

Land management and Urban Form

Very often the research questions which academics and researchers can answer are constrained by the availability of data. Land management data and data from the urban environment can often be difficult to access for academics and researchers. OpenStreetMap offers ready access to a crowdsourced global geodatabase not necessarily limited to the the traditionally popular thematic areas of infrastructure, networks and natural features such as rivers and lakes. Should OpenStreetMap make concerted efforts to develop it’s landcover and landuse capabilities for example as indicated by authors such as Jokar Arsanjani et al (2013b), Estima and Painho (2013) and Estima and Painho (this volume)? If enhanced attribute information about buildings and structures such as height, facade type etc were added to building objects in OpenStreetMap then opportunities for use in urban planning, urban modelling, sustainability and environmental modelling could begin to appear. Are these types of datasets and thematic areas suitable for OpenStreetMap? OpenStreetMap must prevent itself from becoming a collection area for suitably licensed open geographic data just because it fits into OpenStreetMap. This enlargement of the database opens the question of OpenStreetMap as a form of geographic Big Data. Big Data is commonly characterised as being large in volume, produced continuously, and varied in nature. boyd and Crawford (2012:663) argue that "there is little doubt that the quantities of data now available are often quite large, but that is not the defining characteristic of this new data ecosystem". OpenStreetMap contains billions of primitive nodes and ways along with billions of pieces of associated metadata in the form of object attributes, changeset documentation and edit histories. As data and metadata from other thematic areas are added to the OpenStreetMap database the maintenance and long-term sustainability of these additional data must be carefully considered. Extracting knowledge and information from OpenStreetMap as it grows in size may require approaches currently being used or developed from the Big Data domain.

Finally in our outlook we consider how the academic and research community can collaborate with the OpenStreetMap community. The OpenStreetMap community should not be expected to shoulder the entire responsibility for the future of the project. There are roles which the academic community can play in the future development of OpenStreetMap. Presently there are few direct connections between these two communities. There are some academics and researchers who are active 'OpenStreetMappers' but with only empirical evidence we believe these numbers are small. We urge academics and researchers to become actively involved and engaged with the OpenStreetMap community. When academics use data or software purchased from commercial companies we are not slow to report our successes while at the same time reporting bugs, errors and other issues. This model has also worked incredibly well for the Open Source Software community within which the academic and research community are very heavily involved particularly in areas such as operating system design, software engineering and scientific computing. Why should this be any different for OpenStreetMap? It is crucial that the academic community feedback results, observations and recommendations to the OpenStreetMap community from their research. There are numerous ways of doing this. Academic papers and edited volumes such as this one are one such channel. However we urge academics to also interact with the OpenStreetMap community through mailing lists, wikis, social media, etc. Together, in this fashion, everyone will benefit. In a wider sense tantalisingly Graham and Shelton (2013:259) remark that "the futures of geography and big data are still to be made". OpenStreetMap and VGI will have a central role to play in shaping these futures. The future begins here.

Additional References

Agay-Shay, K., Peled, A., Valentin Crespo, A., Peretz, C., Amitai, Y., Linn, S. Nieuwenhuijsen, M. J. (2014). Green spaces and adverse pregnancy outcomes. Occupational and Environmental Medicine, 71(8), 562–569.

Dodge, M. & Kitchin, R. Crowdsourced cartography: mapping experience and knowledge. Environment and Planning A 45, 19 – 36 (2013).

Estima, J., & Painho, M. (2013). Exploratory Analysis of OpenStreetMap for Land Use Classification. In Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information (pp. 39–46). New York, NY, USA: ACM.

Fogue, M., Garrido, P., Martinez, F. J., Cano, J.-C., Calafate, C. T., & Manzoni, P. (2013). An Adaptive System Based on Roadmap Profiling to Enhance Warning Message Dissemination in VANETs. IEEE/ACM Trans. Netw., 21(3), 883–895.

Graham, M. & Shelton, T. Geography and the future of big data, big data and the future of geography. Dialogues in Human Geography 3, 255–261 (2013).

Mooney, P. & Corcoran, P. Analysis of Interaction and Co-editing Patterns amongst OpenStreetMap Contributors. *Transactions in GIS* Volume 18, Issue 5, pages 633–659 (2013b).

Poteete, A. R. & Ostrom, E. Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management. Development and Change 35, 435–461 (2004).

Schellekens, J., Brolsma, R. J., Dahm, R. J., Donchyts, G. V., & Winsemius, H. C. (2014). Rapid setup of hydrological and hydraulic models using OpenStreetMap and the SRTM derived digital elevation model. Environmental Modelling & Software, 61, 98–105.

Stephens, M. Gender and the GeoWeb: divisions in the production of user-generated cartographic information. *GeoJournal* **78,** 981–996 (2013).

Quercia, D. & Saez, D. Mining Urban Deprivation from Foursquare: Implicit Crowdsourcing of City Land Use. *IEEE Pervasive Computing* **13,** 30–36 (2014).