

Has OpenStreetMap a role in Digital Earth Applications?

REVIEW COPY ONLY

Peter Mooney and Pádraig Corcoran

February 26, 2013

Abstract

Volunteered Geographic Information (VGI), as a special subset of crowdsourcing, has provoked interest from many scientific disciplines and industry in the past few years. OpenStreetMap (OSM) is the most famous and well supported example of a VGI project on the Internet today. As a technology, Digital Earth (DE) offers immediate access to enormous quantities of data and information, referenced by locations on the surface of the planet. One of DE's goals is to make global spatial information ubiquitously available as an essential component of democracy. OSM's mission is well positioned within this goal where OSM aims to create a freely available editable map database of the world. This paper explores the characteristics of the collaborative, crowd-based, editing of spatial data in OSM and the crowd itself. We provide some conclusions on how OSM can be best integrated into the fabric of DE.

1 Introduction

Today, in 2013, OpenStreetMap (OSM) is a very famous Volunteered Geographic Information (VGI) project on the Internet with high levels of volunteer participation. VGI, the term coined by Goodchild (2008, 2009), is the collection of spatial data captured by “citizen sensors” where this data is then edited and managed within a collaborative web environment. The unprecedented rates of increase in contribution of spatial data about the world and its environment from citizens to OSM shows no sign of abating. Today OSM has global coverage, is multilingual, is constantly changing and updating, and contains spatial data and attribution representing almost every conceivable geographical feature (Mooney and Corcoran, 2012b). The

ubiquity of the smartphone and Internet-enabled device have seen new digital civilizations forming and harnessing the power of social networking and inter-connectivity. Díaz et al. (2011); Hagenauer and Helbich (2012) both remark that the numbers of people which are both using *and* producing VGI is increasing. Studies have shown that people who produce or contribute to VGI projects tend to focus their attention on areas/regions most familiar to them and this local knowledge is one of the greatest strengths of VGI (De Leeuw et al., 2011).

These characteristics of OSM (and indeed VGI), coupled with the critical mass of hundreds of thousands of volunteer contributors to the project, means that it has the potential to play an important role to play in Digital Earth (DE) applications. Elmes et al. (2009) describe DE as a “virtual, multiresolution, 3-D representation of our planet that enables a person to explore and interact with the vast amounts of natural and social information gathered about the Earth” and acts as an integrated storehouse and source for many forms of geographically referenced information. Elmes et al. predicts DE will manifest itself as an interactive, immersive interface granting users access to a representation of the Earth, which assimilates numerical data, text, images, and maps, along with the relationships between them and models of their underlying processes. From a spatial data perspective DE aims to develop a digital representation of the planet. It is motivated by the need for integrating and interlinking vast geo-referenced, multi-thematic, and multi-perspective knowledge archives which are cross boundary and cross domain (Janowicz and Hitzler, 2012). The purpose of this paper is to discuss the role that OSM can play in DE. Is OSM a suitable candidate as an archive of geographically referenced information? With its strong technological component DE provides a flexible framework to adapt to evolving technologies such as VGI.

1.1 Contributions of this paper

Our invitation to contribute a paper to this special issue of the International Journal of Digital Earth was based on our previous work on OSM and VGI. The editors asked us to consider how OSM would *fit* into Digital Earth if we considered two issues. Firstly, how successful is the collaborative editing process of spatial data in OSM and are there similarities with Wikipedia? Secondly, what characteristics does “the crowd” bring to the data collected and managed within OSM? This paper will discuss these issues by assimilating results from published research on OSM and VGI over the past number of years. Our paper will conclude with a summary of what we believe the

key challenges OSM must overcome in order for it to have consistently high quality data, predictable update cycles, and near homogeneous spatial coverage. The paper is organized as follows. In section 2 we provide an overview of the OSM project. This section is important as it describes how the crowd contributes to OSM and how collaboration occurs. Section 3 is the central section of the paper. In this section we discuss the characteristics of collaborative editing of spatial data in OSM. In particular we draw on research into Wikipedia for similarities. We also discuss the types of contributors within “the crowd” who work on collecting, editing, and managing OSM data. The final section of the paper is Section 5 where we provide some conclusions and outline the key challenges for OSM integration into future DE initiatives. In essence this provides a suggested research agenda for OSM going forward.

2 Understanding how OpenStreetMap works

OpenStreetMap (OSM) was founded in 2004 and has quickly developed into one of the largest, and most well known, sources of VGI on the Internet today. There are several high profile international companies who are now using, or providing products for, OSM data. These include: FourSquare (Foursquare, 2012), Nestoria Property Search (Nestoria, 2011), and MapQuest Navigation and Directions (MapQuest, 2012). World leading GIS company ESRI have produced an add-on component for their famous ArcGIS suite for editing OSM data (ESRI, 2012). Apple began using OSM data in their iPhoto application service in spring of 2012. Wikipedia uses OSM data to render custom maps used by the articles. Flickr uses OSM when their commercial map provider Nokia does not have data available for various cities or regions. The mission of OSM is to create a free database of geographic information for the entire world. Any user who signs up to the OSM project can contribute spatial data to the project or edit data already existing in the database. Within the past few years OSM membership has rapidly increased from a few hundred in mid-2004 to more than 700,000 registered members in November 2012 (Neis et al., 2012). It should be stressed that not all of these registered members are frequently contributing to OSM or actively ‘mapping’ in OSM. Neis and Zipf (2012) show that of these 700,000 registered members almost 200,000 members made at least one edit and that roughly 3% of all members made at least one change per month to the database by the end of 2011. Those members are considered as regular contributors to OSM. The web site *altogetherlost* (altogetherlost, 2012)

provides daily updates of membership statistics and their corresponding contribution rates to OSM. OSM data is available in a variety of formats (XML, Shapefiles, etc.) from a number of sources on the Internet. Data is directly downloadable from OSM using their API. However for those wishing to download large volumes of data the most popular is probably the GeoFabrik service (Geofabrik, 2012) who provide almost near-real-time downloads of OSM data for countries and selected cities and regions. Other places include Topf (2012), for land and water polygons and coastlines, and American company Cloudmade’s OSM download service (Cloudmade, 2012) who provide a large range of value-added downloads including TomTom POI and Adobe Illustrator files.

2.1 Contributing to OSM

Whilst it is not necessary for this article to outline the editing and contribution process of OSM in step-by-step detail it is necessary to summarize the overall collaborative process of editing. A user can contribute new data to OSM in a number of ways. The most popular is the collection of data using a GPS receiver (or GPS enabled device). This data is then loaded into one of the many freely available OSM editing software tools where it can be edited and have metadata attributes (tagging) added. This data can then be submitted and stored in the OSM database. This classical “on-the-ground” mapping approach in OSM can be supplemented with the use of digital photographs or video of the surveyed area, sketch maps, or text descriptions to enhance the quality of the attributes recorded and certain aspects of the spatial data itself. “Mapping Parties” (OpenStreetMap, 2012b) are often organized for localities where small groups of people meet to map new parts of an area or collaboratively edit and update existing data. There is a social aspect to these gatherings which are usually held at pubs or cafés.

Alternatively one can use the available editor tools to trace geometric shapes (polylines, polygons, points) from aerial imagery (available from sources such as Microsoft Bing) which are directly transformed into spatial objects and inserted into the OSM database. This approach is something of a double-edged sword. On the one hand this allows the generation of large quantities of OSM data (where there is aerial imagery available) by contributors without the need for physical “on-the-ground” mapping. However, this approach means the collation of metadata information (street names, area names, landuse types, etc.) about the traced features is difficult or uncertain unless the contributor is very familiar with the area. Mooney et al. (2010a,b) shows that this tracing can lead to both over and under represen-

tation of features. Finally, one can import freely available geodata (provided its license is compatible with OSM) into the OSM database. This is commonly referred to as “bulk-import”. Whilst bulk-import offers the benefits of large volume import of high resolution spatial data with metadata this method is also controversial. Crucially, a bulk-import performed incorrectly can inadvertently delete or overwrite the work of “on-the-ground” mappers which has caused friction and problems in the past (OSM-Talk-List, 2011). Bulk-import must also consider how to translate metadata attributes to OSM metadata or “tags”, and how to properly convert the geometry, including any cleanup and simplification that might be necessary. There are guidelines on how bulk-imports should be performed and these are available on the OSM Wiki. Most importantly bulk-imports should be discussed with the OSM community *before* they are performed. There have been some very successful bulk-imports into OSM in the past, such as the street and road network of the Netherlands. One of the principal reasons for their success was the involvement of the OSM community (particularly OSM contributors in the locality of the bulk import). Using a collaborative approach these datasets were imported slowly and carefully to ensure that maximum benefit was extracted from the import with minimal ‘damage’ to existing OSM data in that region.

Whilst the major of guidelines and tutorial materials for OSM are available online two books have been published by members of the OSM community. Bennett (2010) and Ramm et al. (2010) give comprehensive overviews about the community, the data models, and the software available to work with OSM. They provide people new to OSM with a manual on how to use OSM geodata in their own projects. The books also explain in detail how the geodata is collected and edited for OSM.

2.2 Data Representation in OSM

The geographic information in the OSM database is represented using: Nodes (equivalent to Points), Ways (the collective term for polylines and polygons), and Relations. Relations contain an ordered list of one or more nodes and/or ways as members which are used to define logical or geographic relationships between other elements. Relations are used to model logical (and usually local) or geographic relationships between objects. *Tags*, consisting of a ‘Key’ and a ‘Value’ can be associated with Nodes, Ways, and Relations. Both the key and value are free format text fields. However, in practice there are agreed conventions of how tags are used for most common purposes and these are documented on the “Map Features” page of the

OSM Wiki (OpenStreetMap, 2012a). Anonymous changes to data in the OSM database are not permitted. Editing of the data in the OSM database is predominantly performed using one of the editing software tools available. A registered contributor can change (update, delete) the tags or tag values of any object in the database and can edit the geometry of that object. All changes are timestamped and recorded in the OSM database. Following on from work by other researchers Neis and Zipf (2012) show that whilst there is now over half a million registered contributors only 5% (about 24,000) of these “actively contribute to the OSM project in a more productive way (frequent editing and data contributions)”.

3 Characteristics of Crowdsourced Spatial Data in OSM

As discussed in earlier sections of this paper one of the most unique characteristics of OSM is that “the crowd” collect, edit, and maintain a global database of spatial data. Whilst no concrete statistics are, as yet, available, it is very likely that a large majority of those who contribute to OSM do not have formal professional training in surveying, cartography, or GIS. These types of skills are a prerequisite for working with commercial or governmental entities producing spatial data and mapping products. In this section we discuss characteristics of the spatial data collected by OSM contributors. We address this under three headings: geometric accuracy of the collected data, attribution and semantic accuracy, and general characteristics of “the crowd”. Hagenauer and Helbich (2012) argue that discussions about quality and completeness of VGI data is a complex, non-trivial task and involves dealing with a large community of contributors whose motivations for participating, contributing and using spatial data can differ substantially.

3.1 Collaborative Editing in OSM

In this section we discuss the practicalities of collaborative editing in OSM. OSM provides a Wiki-based model for the collection, editing, and updating of geospatial data collected by volunteers. Any OSM contributor can edit and update the spatial data (or annotation information) supplied by any other OSM contributor. Matei and Dobrescu (2010) explains that while philosophically akin with, although not completely similar to, the open-source software movement, Wikipedia and OSM rely on the wiki publishing paradigm. Wikis are Web-based collective and non-hierarchical publication

systems with Web-based editing interfaces which allow any Internet visitor to add, delete, and publish content. The history and discussion pages of Wikipedia often provide an insightful story into the varied and often conflicting contributions and edits performed on an article or group of articles. Unfortunately, few readers actually investigate these pages with most readers naturally more interested in the article page itself. In OSM the history of objects and/or the changesets are available for viewing by users of OSM data in order to better understand the lineage of the data. However as Mooney and Corcoran (2012b) and Neis et al. (2011) point out the vast majority of users are interested in the actual data and not how the OSM database ‘evolved’ to this point. Currently, there are only a few examples in the literature available investigating the nature of collaborative editing in OSM. Specifically we are referring to analysis of how many contributors are working in a specific area/region and what are the characteristics of their data and metadata contributions to OSM in general and are there specific strata of contributors to OSM based on their contributions. Neis and Zipf (2012); Mooney and Corcoran (2012d) both show that the trend in OSM is towards a small subset of contributors performing a large percentage of the overall editing work. Mooney and Corcoran (2012d) show for the entire OSM history of Berlin, London, and Paris the 10 most prolific contributors perform approximately 45% of all edits. The top 10% of contributors in those cities are responsible for over 80% of all edits. Neis and Zipf (2012) analyse the entire history of OSM and find that globally major contributors, or “senior mappers”, represent only 5% of all OSM members. Incredibly almost 62% of OSM members (about 300,000) have never performed a single edit. The combined work of Mooney and Corcoran (2012c) and Mooney and Corcoran (2012d) investigate if collaboration is actually taking place in OSM data editing. They extract collaboration links between contributors from an analysis of objects that two or more contributors have “co-edited”. They conclude from their analysis, and without the benefit of explicit OSM information on collaboration/friendships, that any collaboration is almost purely incidental. With the exception of a small number of people most contributors appear to be contributing data in isolation and restricting editing to their own data. Suh et al. (2009) indicates that approximately 68% of edits in Wikipedia are actually restoration and reverts of edits to page content. In the English Wikipedia the contribution of knowledge and information is similar to the results reported for OSM. Approximately 2% of all Wikipedia contributors have made more than 100 edits (see WikipediaStats (2012)). In the next section we will compare collaborative editing practices in OSM and Wikipedia.

3.2 Collaborative Editing in OSM - comparison with Wikipedia

Is the crowdsourced collaborative process of editing and managing spatial data in OSM similar to collaborative knowledge management in Wikipedia? In table 2 we provide a brief comparison of OSM and Wikipedia under a number of headings including: dispute resolution, administration, and verifiability of data. It is clear from table 2 that there are many similarities and differences between OSM and Wikipedia in their roles as wiki-based knowledge management projects. OSM is relatively small in comparison to Wikipedia. The contents of OSM (ways only) range from simple 4 node polygons representing garden sheds to polylines representing continental coastlines. In a similar way Wikipedia articles range from trivial to long, heavily researched and edited, articles on science, politics, and arts. Verifiability and dispute detection/resolution are very similar with both projects employing a mixture of automated (via bots) and human surveillance. In both OSM and Wikipedia anonymous edits are not permitted. All edits and reverts are carefully timestamped and logged.

From a management viewpoint one must be careful with the context within which comparisons of OSM and Wikipedia are performed. The management structure and IT infrastructures behind Wikipedia completely dwarf OSM. Alongside the Wikipedia community of volunteers there is the Wikimedia Foundation (WMF). WMF has about 100 employees and in the Q3 and Q4 of 2011 WMF had an income of US 27 million (Wikimedia, 2012). In terms of management of the actual pages on Wikipedia, Geiger (2011) points out that in the English language Wikipedia contributors who have made at least 1,000 edits to the project in a given year are eligible to run for election to the prestigious Arbitration committee for Wikipedia. This committee makes decisions on disputes which have arisen and where there are no clear community consensus on an issue. This arbitration is necessary in the absence of strict peer-review structures because as Matei and Dobrescu (2010) remarks that once Wikipedia “abandoned the peer-review component and became a *free-for-all* collaborative space” it became “an instant success”.

In OSM the OSM Foundation (OSMF) is currently a membership organization where all members decide who gets on the board. The OSMF does not employ any staff. OSMF has a total annual income of about £100,000. With no paid staff, the OSMF board members do all of the Foundation work themselves, or find other volunteers to do it. This is what makes OSM all the more remarkable as a VGI project. Volunteers provide administration for servers, write code for website and API services, provide data download

services, and act as moderators or gatekeepers for data related to their local area. The IT infrastructure is modest in comparison to that of Wikipedia. OSM rely on donations from the community to purchase hardware or donations of server space and hosting facilities. The OSM Data Working Group (OSM-DWG) are akin to the Arbitration committee in Wikipedia. This group of volunteers deal with resolution of issues in copyright violation, disputes, vandalism, and bots, and issues which are beyond resolution by normal means in the community. With community support they help set policy on OSM data. The OSM-DWG are usually asked to make decisions regarding data vandalism and on bulk imports which do not comply with community guidelines. The OSM-DWG is small with about 7 members. They prefer to stand back and allow the local OSM community deal with local disputes or problems. In the next section we attempt to profile the community members who volunteer for OSM.

3.3 Crowd Characteristics

VGI projects are more popular in urban areas because of the larger population of people living there in comparison to more rural locations. There exists very detailed coverage of towns and cities in OSM but there is less knowledge of rural areas to the point where geometric issues become less important than attribution and metadata (Mooney and Corcoran, 2012a). Naturally crowds are larger in urban areas but this can lead to an inhomogeneous distribution of contribution effort and map representation. Rural areas are generally less well mapped in OSM than their urban counterparts. As Mooney et al. (2010a,b) explain that while road and street networks are very well mapped there is dramatically less interest in land cover mapping. Figure 1 shows an example of an urban area, near Nottingham UK, where beyond the extent of rural development there is little or no land-cover representation. Land-cover mapping is difficult as it requires mappers to physically access the countryside and terrain. Tracing directly from aerial imagery would, in most cases, not be recommended.

There are also different types of crowd members. Contributors to VGI projects can have deep local knowledge, those who prefer to physically survey areas and contribute the results, conscientious members, etc. Individual people have different motivations for doing things, and those motivations create different levels of participation in VGI projects. Most participants will start as consumers, and only a small percentage of these will eventually move to contribute, collaborate, and act as meta-designers, and thereby be responsible for the content that is shared with everyone. Understanding why

this occurs is what Fischer (2011) calls the “startup paradox”. As discussed earlier Neis and Zipf (2012) found that only 3–5% of all registered members of OSM “actively contribute to the project in a productive way” and Mooney and Corcoran (2012d) showed specific examples such as the situation where 10 contributors were responsible for 45% of all edits in London, Paris, and Berlin. Perkins and Dodge (2008) provide a detailed overview of a mapping party in OSM involving a crowd of mappers. Whilst their mapping party in Manchester UK attached a reasonably inhomogeneous set of people Perkins and Dodge (2008) are concerned that the motivation behaviour for some OSM contributors “seems to conform to particular male stereotypes, with a strongly competitive ethos revealed in online posting of league tables of the numbers of streets uploaded or tagged and a desire to control their part of the map”.

It is very difficult to predict how a self organizing community such as the community of contributors in OSM could possess the resources to actually provide homogeneous representation of all areas in a region containing large urban and rural areas. Potentially OSM should focus on ‘where the crowd is’ leaving homogeneous spatial coverage to commercial mapping companies of National Mapping and Cadastral Agencies (NMCA) who have the financial and human resources to complete these tasks. In developed countries it is very often the case that OSM is ‘very like’ the NMCA where it provides very high resolution spatial data for urban areas. The update and refresh rates of data for these areas are often very quick in OSM. However it appears that this process is reversed in many developing countries where citizen-based mapping efforts are often far superior to the mapping products generated by the NMCA in those countries (De Leeuw et al., 2011; Richter et al., 2010; Zook et al., 2010).

3.4 Motivations of Contributors to VGI

Coleman et al. (2009) attempt to understand the motivations of the crowd by investigating the open source software development world and Wikipedia amongst others. They emphasise that government agencies and academia must learn more about the motivations and social expectations in the VGI community before the data and information from VGI is dismissed. Can the contributions of “the crowd” be relied on for the long-term? In his report McLaren (2011) asks that “with so many crowdsourced projects contending for the attention of the citizen, will fatigue and lack of interest over time make citizen contributions a scarce resource”. This raises the question of possible need for incentives for people to contribute to VGI projects. OSM

is changing all of the time. It does not adhere to the strict update and release cycles of commercial spatial data providers or NMAs. Its collaborative model means that there might often be areas of conflict, damage to data, mistakes or inconsistencies introduced. But this is captured well by Dodge and Kitchin (2011) who call OSM a “live laboratory of shifting ontological politics in cartography” and that it might well be the case the OSM and VGI in general will develop in symbiosis with professional mapping and cartography. In professional and commercial mapping and cartography quality control is of vital importance. It is very difficult to extract data and information, in an automated way, to make judgements or assessments about what motivates contributors to work for VGI projects. Budhathoki et al. (2010) is an important contribution our understanding of motivation in VGI. Budhathoki et al. argues that while VGI is usually “created in collaborations with users who usually do not have special skills in handling spatial data trained people and professionals also play major role in contribution in that VGI”. Lin (2011) carried out interviews of OSM activists and contributors at an OSM conference in 2010 to attempt to better understand these motivations. Lin sees OSM as “a boundary object that enables actors from different social worlds to co-produce a map through interacting with each other and negotiating the meanings of mapping, the mapping data and the map itself”. The results of Lin (2011)’s study was closely aligned to the Budhathoki et al. conceptual framework about motivation of contributors of VGI where motivation grows from a mixture of personal, social, technological, intrinsic, extrinsic reasons. Budhathoki et al. used conversational text among Openstreetmap Users (talk pages) and surveys with open-ended questions and used grounded theory to define the motivation behind contributors work in VGI.

In the next section we review results on the geometric accuracy of the spatial data managed by the OSM project.

3.5 Geometric Accuracy of OSM data

Geometric accuracy of crowdsourced spatial data is often cited as one of the biggest areas of concern regarding more widespread use of VGI. Does OSM have geometrically accurate data? Can a very loosely coordinated crowd contribute, and manage, an accurate database of spatial data for the world? This GIS community has tended towards what Goodchild and Li (2012) call a “geographic approach” to quality assessment in VGI as “it is attractive to geographers because it taps the heart of geographic knowledge”. Many academic studies (Haklay et al., 2010; Koukoletsos et al., 2012; Girres and

Touya, 2010; Ludwig et al., 2011; Neis et al., 2011; Over et al., 2010; Haklay, 2010) have investigated VGI with specific emphasis on OSM. Overwhelmingly the results of these studies have shown that the geometric accuracy of OSM is very high and in some cases even surpasses that of National Mapping Agency datasets. In this section we briefly discuss some of these studies.

Neis et al. (2011) comment that in the past OSM accuracy was good enough for the production of maps. But now in countries where OSM is mature and well developed “it is on a par with commercial producers of spatial data both spatially and temporally”. Neis et al.’s studies showed that in 2010 the German OSM provided 27% more data (road networks) than proprietary sources at that time. In a similar study in the United States Zielstra and Hochmair (2011b) showed that there were significant differences were coverage of rural and urban areas but overall comparison of OSM data to proprietary data in US cities were very similar for all data providers. Geometric accuracy in OSM has been shown to improve as more contributors become involved in the editing process. Haklay et al. (2010) used the guidance of empirical evidence from open source software development to show improvements in accuracy as more contributors became involved. Girres and Touya (2010) showed similar examples for OSM in France while Neis et al. (2011) and Ludwig et al. (2011) performed this analysis in Germany. One of the shortfalls of studies of geometric accuracy in OSM is that these comparisons are usually performed for OSM in countries where there are flexible mechanisms for accessing authoritative, or gold standard, datasets or where researchers have been given special access to commercial data. As Goodchild and Li (2012) point out it is often expensive and time-consuming to assemble the reference data needed to compare VGI against. In the United States there were serious issues in the OSM road and street network resulting from a poorly orchestrated import of freely available TIGER data. However these problems are being rectified quickly. Little or no work has been carried out on the geometric accuracy of landuse objects or coastlines. Mooney et al. (2010b) compared lakes/waterways in OSM to NMA data. They found that representation of natural features was poor for OSM in areas of Ireland where it was physically difficult to collect primary survey data (for example mountain regions).

Some work has been reported on using a crowdsourced VGI approach to landcover validation and associated data gathering. Fritz et al. (2009) introduced the Geo-Wiki Project which is a global network of volunteers who are trying to improve the quality of global land cover maps. They have since extended this work and reported the results in Fritz et al. (2012). In Fritz et al. (2009) volunteers are asked to review hotspot maps of global

land cover disagreement and determine, based on what they actually see in Google Earth and their local knowledge, if the land cover maps are correct or incorrect. Their input is recorded in a database, along with uploaded photos. This database can then be used in the future for land cover validation or potentially the creation of a new and improved hybrid global land cover map. In Fritz et al. (2012) they extend the discussion of Geo-Wiki. They stress that the “proposed tool does not intend to replace current land cover validation activities by experts, but can potentially complement some of those activities”. Geo-Wiki can help provide either additional data or software tools to undertake validation. In relation to the discussion in Section 3.3 the authors call Geo-Wiki “more of an expert-sourcing system than a crowdsourcing one” and emphasise that to reach out to a wider crowd and build a sustainable community they will need to integrate additional social networking tools and feedback mechanisms that motivate individuals to participate. In the final part of this section we will discuss issues relating to metadata in OSM.

3.6 Metadata Issues in OSM

One of the easiest tasks in VGI is to “tag” objects (data objects, photographs, videos, etc.) with text attributes describing the characteristics of those objects. In OSM this is performed using key-value pairs called “tags”. OSM provides an extensive list of the most commonly used tags but contributors are free to make up their own tags. Whereas OSM has been shown to be strong in terms of geometric accuracy there are issues around attribute accuracy. Ballatore and Bertolotto (2011) call OSM “semantically poor” stemming from the fact that while the ontology for tagging is extensive it is often used incorrectly, or not at all. Al-Bakri and Fairbairn (2012) found in their study of OSM and OS-GB data that there was major heterogeneity between authoritative data and VGI data which they stress is “a major issue”. They recommended more detailed documentation in VGI regarding how data and schema are created and maintained. Studies such as Mooney and Corcoran (2012a), Al-Bakri and Fairbairn (2012), and Goodchild and Li (2012) show OSM data has semantic characteristics which make quality assessment more difficult due to the possible presence of local bias, semantic inconsistencies, etc. Poore and Wolf (2012) introduce their term “Metadata squared” which describes how loosely the term metadata relates to VGI. Poore and Wolf argue that VGI data has become metadata and vice versa. But beyond the produced map is a “vast, swirling universe of other data that describe the map, consisting of computer programs, tiling schemes,

web pages, IRC chat rooms, YouTube video demonstrations, tweets, discussion groups, and wiki pages. All of these are socially mediated, produced by a community, and accessible to any user.” Making sense of this information is a difficult task. Attribution accuracy and management is not an easy task in VGI. Richter and Winter (2011) concludes an overview of future possibilities for VGI by arguing that “the management of user-added semantics is still in it’s infancy” in databases today. Mooney and Corcoran (2012a) and Mooney and Corcoran (2012b) show examples of “tag flip-flopping”. The authors believe that these “tag disputes” or “tag wars” indicate that different contributors have different ideas of what values should be assigned to tags. Amongst the tags which display the highest frequency of changes include the “name” attribute for objects such as streets and the designation of roads in the “highway” tag. In Figure 1 a summary of the number of times values assigned to the “highway” tag to indicate the designation (class) of roads is shown. The summary provides results for frequently edited objects in four countries. Given the open nature of tagging in VGI and OSM there is a high possibility of incorrect or spurious values being assigned as tag values. In the case of intense tag disputes the OSM Data Working Group can help to resolve the issue. Automated approaches are more complex to develop. Richter and Winter believe that to detect, and fix, problems with user-added semantics will require a mixed approach of human intelligence input and data mining approaches. Coleman et al. (2009) argue that industry and government will have to “accept the fact that VGI is a perpetually unfinished artefact” and tools for filtering and fusing this information with other sources are required.

Highway	UK	Ireland	Austria	Germany
1	4,999(59.4%)	298(50.5%)	1110(47.1%)	495(54.8%)
2	2,621(31.2%)	222(37.6%)	855(36.3%)	271(30.0%)
3	650(7.7%)	60(10.2%)	305(12.9%)	110(12.2%)
4	117(1.4%)	8(1.4%)	78(3.3%)	22(2.4%)
≥ 5	22(0.3%)	2(0.3%)	10(0.4%)	5(0.6%)

Table 1: A summary of “tag flip-flopping” from Mooney and Corcoran (2012b). The table shows the number of times the value of the “highway” tag was changed for a large subset of highway objects in four countries with good OpenStreetMap coverage

4 Issues for OSM in Digital Earth

In this section we shall discuss the key challenges for OSM integration into DE. Grossner et al. (2008) conclude that in 2008 they found that DE “presently comprises a nascent software development platform, multiple application software programs and a loosely organized intellectual movement”. Guo et al. (2010) remark that DE is now finally becoming a reality and “provides a fusion base for multisource geospatial data”. This integrated processing and fusion of different kinds of data and models in the same environment “will certainly have new research needs”. Following on from this Annoni et al. (2011) remark that from a European perspective on the development of DE there are emerging needs for “exploiting the potential of human-centric sensing”. We feel that DE has the opportunity to benefit from two discrete technological worlds: the formal institutional Spatial Data Infrastructures (SDI) “top-down” world and the VGI “bottom-up” world. In many senses OSM is itself a DE application which can be used by other applications, models, and services. Wen et al. (2012) argues this case for geographical modelling where there must be “suitable infrastructures to support the open sharing and effective execution of models”. OSM has demonstrated its ability to provide support for open sharing and collaborative editing of spatial data. Scale is an ongoing issue. In the future scenario of increasing contributions, greater global coverage, and increased usage by commercial companies and services OSM must scale to meet these growing demands. To play a successful role in DE we feel that OSM will need to scale both technologically and in terms of the quality assurance mechanisms that are put in place. As discussed in Section 3.2 OSM’s organisational structure and technical infrastructure is currently dwarfed by that of Wikipedia. The heterogeneity found in some areas of OSM such as tagging of features or the resolution at which geographical features are sampled and subsequently modelled is an ongoing problem but are being addressed by researchers. Janowicz et al. (0) emphasises that the vision of DE calls for “stronger integration of dynamic information systems (for example: sensor observations) and new sources of information (for example: VGI)”. The integration of sensor webs and VGI has meant that ontological models have gained greater importance. Heterogeneity is a major obstacle towards this integration an example of which was shown for OSM in Table 1. There are some examples where the heterogeneity of semantics in VGI are tackled in semi-automatic workflows with Craglia et al. (2012) developing a successful software system for forest fire detection using large volumes of unstructured data from social media (Twitter and Flickr). Janowicz and Hitzler (2012) stresses that

VGI is created and maintained by a highly-heterogeneous user community with different backgrounds and application areas in mind. VGI relaxes the typical rules under which data is collected for the benefit of providing the most up-to-date data. Janowicz and Hitzler does not see this as a negative aspect of VGI and rather they believe that “this source of variety opens up new possibilities for science and especially for the evaluation of data in DE”. Both OSM and DE are growing and evolving. We feel that the answers to research questions from one will benefit the other. The next section is the final section of the paper and provides some overall conclusions and future outlook.

5 Conclusions and Outlook

The drivers of VGI are billions of what Goodchild (2008) refers to as “potential citizen sensors” collecting information in both explicit and passive modes and are distributed across the globe. Currently, for OSM this distribution is skewed towards urban areas with higher rates of “buy-in” to OSM in certain countries. OSM is a VGI project where a more pro-active, higher-work load, participation level is needed. At time of writing there are approximately 700,000 registered members of OSM. About 35% of these perform only one edit with only 10% doing what Neis and Zipf (2012) calls “editing of consequence”. However, these contributors have produced an enormous body of work in OSM and this spatial data has “good attribute quality and high geometric accuracy” (Haklay, 2010). This paper has provided a detailed overview of OSM with the goal of understanding if OSM can be a successful component of DE. The paper has delivered a number of important contributions. Firstly, we have provided a detailed overview of the current state-of-the-art literature on VGI with specific focus on OSM. Secondly, we have provided discussion of OSM under a number of key headings: collaborative creation of spatial data, the characteristics of “the crowd”, and geometric and semantic accuracy of the produced spatial data. We have also, in section 4, discussed the issues related to OSM integration into DE.

As discussed in section 4 there are some issues in OSM which must be addressed for future research. Some researchers have moved beyond these issues to conducting *value-added* research into OSM: Jiang and Liu (2012) for automated delineation of urban areas, Mooney and Corcoran (2012d) on the social interaction amongst OSM contributors, Goetz (2012) on the development of 3-D indoor routing applications, urban land-use patterns on the basis of OSM and the delineation of continuous urban areas (Hagenauer and

Helbich, 2012), public transit accessibility (Zielstra and Hochmair, 2011a), enrichment of 3D city models with VGI data (Smart et al., 2011), and Corcoran and Mooney (2012) who study street network evolution. Given our experience of research in this area and the review of the literature we believe that there are two major areas where OSM can continue to improve and subsequently improve its integration into DE. These two areas are increasing the general usage of VGI and OSM and the sustainability of crowdsourcing spatial data into the future. The next two paragraphs provides an overview discussion of these areas.

1: Increasing usage of VGI and OSM:

In essence VGI appears to have all of the required ICT ingredients to provide a dynamic picture of the global environment for DE. VGI can leverage large numbers of dedicated “citizen sensors” to collect enormous amount of spatial data. Citizens, experts and non-experts alike, are increasingly participating in the process of generating continuous spatial information and collaborating with others in problem-solving tasks. Ho and Rajabifard (2010) argue that as a visible representation of citizen sentiment VGI can be a “potential barometer for peoples environmental concerns and attitudes and potentially lead to better citizen buy-in to SDIs”. The fact remains that VGI, as a spatial form of the user-generated content in Web 2.0, raises serious concerns and reservations within the GIS, Geomatics, and Environmental Science communities regarding its quality, authenticity, and sustainability. Despite studies highlighting OSM’s high quality compared to NMA data and other very positive results outlined earlier in this paper, VGI and OSM lack usage “in serious geomatics applications” (Over et al., 2010). We feel that some large flagship applications or commercial products adopting OSM as a **primary source of spatial data** are urgently required. DE could provide an application platform. OSM involvement or integration in a regional or national SDI is another possibility. Despite concerns surrounding quality and authenticity we believe that inhomogeneity in terms of coverage in OSM is its most serious challenge and prevents adoption of OSM as a spatial data provider. How can crowds be motivated to collect data in: rural areas, Third World countries, socially-disadvantaged urban areas, etc? Hagenauer and Helbich (2012) conclude that OSM is not yet a replacement for professionally acquired data sets due to these very issues.

2: Sustainability of crowdsourcing and VGI:

OSM integration into DE raises concerns about the sustainability of crowdsourcing. Will “the crowd” continue to contribute to OSM? How much more work can be feasibly expected from those top contributors to OSM in the face of: criticism of OSM from industry and academia, bulk imports deleting work done by real “on-the-ground” mappers, potentially diminishing personal rewards from contributing to the OSM project, or the possibility that there is “no-more interesting work” left to do. Dedicated and productive members who actively contribute to community efforts are crucial to the success of collaborative wiki-based projects. Prasarnphanich and Wagner (2011) state that “to have collaboration and sustainable ecosystems emerge from wiki-type technologies is not obvious and requires justification”. Wang et al. (2012) suggest that the social aspect of crowdsourcing knowledge products could be sustained if individuals develop social networks within the project which reduces their likelihood of withdrawing from the project despite not having strong bonds with the project itself. So is Wikipedia sustainable? Prasarnphanich and Wagner (2011) argues that it is based on a critical mass of ‘Core’ contributors (very frequent contributors). They discuss empirical evidence that indicates that this ‘Core’ group is transient but new ‘Core’ groups constantly emerge from contributors who previously had more peripheral contribution levels.

From a spatial data perspective it is not guaranteed that a certain object or group of objects will ever be mapped. OSM’s reliance on volunteered contributions inherently means some places and objects will be overlooked. Bulk imports may not be available, or suitable, for these areas. Globally the problem of the “digital divide” can be a contributing factor for incomplete mapping of less well developed countries (Hagenauer and Helbich, 2012; De Leeuw et al., 2011; Haklay et al., 2010; Zook et al., 2010; Richter et al., 2010). Local knowledge is a vital ingredient of OSM. As high frequency OSM contributors finish mapping their own areas there is scope for them to assist mapping other areas as happened during the Haiti and Sendai earthquakes. Lam and Riedl (2011) remarks that as more contributors arrive at a wiki-system individuals will find it increasingly difficult to contribute without coordinating their efforts with others. OSM may never scale to having local mappers *everywhere*. However, as Prasarnphanich and Wagner (2011) points out, the key to sustainability of projects such as Wikipedia and OSM are those “small unnoticeable but countable” tasks or edits which are crucial in maintaining the quality of the information without necessarily expanding the coverage. Richter and Winter (2011) believe that long-term VGI sustain-

ability will be achieved if citizens can contribute content to projects where “these contributions are facilitated unobtrusively, casually, or, even calmly”. Jones and Weber (2012) agree and suggest that the current one-size-fits all software interfaces for crowdsourcing spatial data may alienate new inexperienced contributors and expert veteran contributors in equal measure. We feel that these types of considerations extend to all DE applications. Janowicz and Hitzler (2012) embraces the concept of VGI for DE stressing that using VGI will make DE “call for methods to reason in the presence of heterogeneous and contradicting conceptual models, while maintaining the variety brought in by different scientific domains”.

We believe that OSM will continue it’s growth for the foreseeable future. It is a project requiring careful examination of the underlying social and cultural processes in the gathering, submission, management, distribution, and usage of VGI. By it’s very nature: global coverage, a large number of “citizen sensors”, etc. OSM has a role to play in DE. In VGI there are issues regarding quality, verifiability, and sustainability and these offer a very “well-motivated research problem” (Goodchild and Li, 2012). But as Cuff et al. (2008) argue citizen interest in VGI and citizen science will continue to grow and what is considered “today’s exotic and disturbing data collection practices will appear banal in 10 years hence”. The future research topics, as suggested, for OSM in DE are numerous and these unanswered questions will provide substantial work for researchers in the coming years.

References

- Al-Bakri, M. and Fairbairn, D. (2012). Assessing similarity matching for possible integration of feature classifications of geospatial data from official and informal sources. *International Journal of Geographical Information Science*, 0(0):1–20.
- altogetherlost (2012). Daily updated statistics about openstreetmap. Online: <http://osmstats.altogetherlost.com/>.
- Annoni, A., Craglia, M., Ehlers, M., Georgiadou, Y., Giacomelli, A., Konecny, M., Ostlaender, N., Remetey-Fülöpp, G., Rhind, D., Smits, P., and Schade, S. (2011). A european perspective on digital earth. *International Journal of Digital Earth*, 4(4):271–284.
- Ballatore, A. and Bertolotto, M. (2011). Semantically enriching vgi in support of implicit feedback analysis. In Tanaka, K., Fröhlich, P., and Kim,

- K.-S., editors, *Web and Wireless Geographical Information Systems*, volume 6574 of *Lecture Notes in Computer Science*, pages 78–93. Springer Berlin / Heidelberg.
- Bennett, J. (2010). *OpenStreetMap: Be your own cartographer*. Packt Publishing Ltd, Olton Birmingham, B27 6PA, UK.
- Budhathoki, N., Nedovic-Budic, Z., and Bruce, B. (2010). An interdisciplinary framework for understanding volunteered geographic information. *Geomatica*, 64(1):11–26.
- Cloudmade (2012). Openstreetmap data downloads. Online: <http://downloads.cloudmade.com/>.
- Coleman, D. J., Georgiadou, Y., and Labonte, J. (2009). Volunteered geographic information: The nature and motivation of producers. *International Journal of Spatial Data Infrastructures Research*, 4(1):332–358.
- Corcoran, P. and Mooney, P. (2012). Characterising the metric and topological evolution of openstreetmap street network representations. *European Physical Journal Special Topics (special issue Spatially Embedded Socio-Technical Complex Networks)*, page to appear.
- Craglia, M., Ostermann, F., and Spinsanti, L. (2012). Digital earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, 5(5):398–416.
- Cuff, D., Hansen, M., and Kang, J. (2008). Urban sensing: Out of the woods. *Communications of the ACM*, 51(3).
- De Leeuw, J., Said, M., Ortegah, L., Nagda, S., Georgiadou, Y., and De Blois, M. (2011). An assessment of the accuracy of volunteered road map production in Western Kenya. *Remote Sensing*, 3(2):247–256.
- Díaz, L., Granell, C., Gould, M., and Huerta, J. (2011). Managing user-generated information in geospatial cyberinfrastructures. *Future Generation Computer Systems*, 27(3):304–314.
- Dodge, M. and Kitchin, R. (2011). Mapping experience: Crowdsourced cartography. *Social Sciences Research Network*.
- Elmes, G., Weiner, D., and D’Alessandro-Scarpari, C. (2009). Digital earth. In Chief: Rob Kitchin, E. and Thrift, N., editors, *International Encyclopedia of Human Geography*, pages 198–204. Elsevier, Oxford, 1st edition.

- ESRI (2012). Arcgis editor for osm server component and/or arcgis editor for osm desktop. Online press release: <http://esriosmeditor.codeplex.com/releases/view/84803>.
- Fischer, G. (2011). Understanding, fostering, and supporting cultures of participation. *ACM Journal of Interactions*, XVIII(3):42–55.
- Foursquare (2012). foursquare is joining the openstreetmap movement! say hi to pretty new maps! Online at the Foursquare official blog: <http://blog.foursquare.com/2012/02/29/foursquare-is-joining-the-openstreetmap-movement-say-hi-to-pretty-new-maps/>.
- Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., and Obersteiner, M. (2009). Geo-wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1(3):345–354.
- Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., van der Velde, M., Kraxner, F., and Obersteiner, M. (2012). Geo-wiki: An online platform for improving global land cover. *Environmental Modelling and Software*, 31(0):110 – 123.
- Geiger, R. S. (2011). The lives of bots. In Lovink, G. and Tkacz, N., editors, *Critical Point of View: A Wikipedia Reader. INC Reader 7*, pages 78–93. Institute of Network Cultures, Amsterdam, The Netherlands.
- Geofabrik (2012). Openstreetmap data downloads. Online: <http://download.geofabrik.de/>.
- Girres, J.-F. and Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*, 14(4):435–459.
- Goetz, M. (2012). Using crowdsourced indoor geodata for the creation of a three-dimensional indoor routing web application. *Future Internet*, 4(2):575–591.
- Goodchild, M. F. (2008). Whither vgi? *GeoJournal*, 6(72):239–244.
- Goodchild, M. F. (2009). Neogeography and the nature of geographic expertise. *Journal of Location Based Services*, 3(2):82–96.
- Goodchild, M. F. and Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1(1):110–120.
- Greenstein, S. (2012). The range of linus’ law. *Micro, IEEE*, 32(1):72.

- Grossner, K. E., Goodchild, M. F., and Clarke, K. C. (2008). Defining a digital earth system. *Transactions in GIS*, 12(1):145–160.
- Guo, H., Liu, Z., and Zhu, L. (2010). Digital earth: decadal experiences and some thoughts. *International Journal of Digital Earth*, 3(1):31–46.
- Hagenauer, J. and Helbich, M. (2012). Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *International Journal of Geographical Information Science*, 26(6):963–982.
- Haklay, M. (2010). How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. *Environment And Planning B: Planning and Design*, 37(4):682–703.
- Haklay, M., Basiouka, S., Antoniou, V., and Ather, A. (2010). How many volunteers does it take to map an area well? the validity of linus’ law to volunteered geographic information. *The Cartographic Journal*, 47(4):315 – 322.
- Ho, S. and Rajabifard, A. (2010). Learning from the crowd: the role of volunteered geographic information in realising a spatially enabled society. In *GSDI 12 World Conference: Realising Spatially Enabled Societies*, Singapore.
- Janowicz, K., Brring, A., Stasch, C., Schade, S., Everding, T., and Llaves, A. (0). A restful proxy and data model for linked sensor data. *International Journal of Digital Earth*, 0(0):1–22.
- Janowicz, K. and Hitzler, P. (2012). The digital earth as knowledge engine. *Semantic Web*, 3(3):213–221.
- Jiang, B. and Liu, X. (2012). Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. *International Journal of Geographical Information Science*, 26(2):215–229.
- Jones, C. E. and Weber, P. (2012). Towards usability engineering for online editors of volunteered geographic information: A perspective on learnability. *Transactions in GIS*, 16(4):523–544.
- Koukoletsos, T., Haklay, M., and Ellul, C. (2012). Assessing data completeness of vgi through an automated matching procedure for linear data. *Transactions in GIS*, 16(4):477–498.

- Lam, S. and Riedl, J. (2011). The past, present, and future of wikipedia. *Computer*, 44(3):87–90.
- Lin, Y.-W. (2011). A qualitative enquiry into openstreetmap making. *New Rev. Hypermedia Multimedia*, 17(1):53–71.
- Ludwig, I., Voss, A., and Krause-Traudes, M. (2011). A comparison of the street networks of navteq and osm in germany. In Geertman, S., Reinhardt, W., Toppen, F., Cartwright, W., Gartner, G., Meng, L., and Peterson, M. P., editors, *Advancing Geoinformation Science for a Changing World*, volume 1 of *Lecture Notes in Geoinformation and Cartography*, pages 65–84. Springer Berlin Heidelberg.
- MapQuest (2012). Mapquest-osm tiles + mapquest open aerial tiles. Online at the Open Mapquest official blog: <http://developer.mapquest.com/web/products/open/map>.
- Matei, S. A. and Dobrescu, C. (2010). Wikipedia’s neutral point of view (npov): Settling conflict through ambiguity. *The Information Society*, 27(1):40–51.
- McLaren, R. (2011). Crowdsourcing support of land administration: A new, collaborative partnership between citizens and land professionals. Technical Report none, Royal Institution of Chartered Surveyors commissioned by RICS Land Group, RICS, Parliament Square, London SW1P 3AD United Kingdom.
- Mooney, P. and Corcoran, P. (2012a). The annotation process in openstreetmap. *Transactions in GIS*, 16(4):561–579.
- Mooney, P. and Corcoran, P. (2012b). Characteristics of heavily edited objects in openstreetmap. *Future Internet*, 4(1):285–305.
- Mooney, P. and Corcoran, P. (2012c). How social is openstreetmap? In *Proceedings of the The 15th AGILE International Conference on Geographic Information Science*, AGILE 2012, pages 514–518. AGILE.
- Mooney, P. and Corcoran, P. (2012d). The role of communities in volunteered geographic information projects. In Krisp, J., editor, *Proceedings of the 9th Symposium on Location Based Services*, volume x of *Lecture Notes in Geoinformation and Cartography*, page (to appear). Springer Berlin / Heidelberg.

- Mooney, P., Corcoran, P., and Winstanley, A. C. (2010a). A study of data representation of natural features in openstreetmap. In *Proceedings of the 6th GIScience International Conference on Geographic Information Science*, GIScience 2010, page p150. University of Zurich.
- Mooney, P., Corcoran, P., and Winstanley, A. C. (2010b). Towards quality metrics for openstreetmap. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '10, pages 514–517, New York, NY, USA. ACM.
- Neis, P., Goetz, M., and Zipf, A. (2012). Towards automatic vandalism detection in openstreetmap. *ISPRS International Journal of Geo-Information*, 1(3):315–332.
- Neis, P., Zielstra, D., and Zipf, A. (2011). The street network evolution of crowdsourced maps: Openstreetmap in Germany from 2007 to 2011. *Future Internet*, 4(1):1–21.
- Neis, P. and Zipf, A. (2012). Analyzing the contributor activity of a volunteered geographic information project - the case of openstreetmap. *ISPRS International Journal of Geo-Information*, 1(1):1–23.
- Nestoria (2011). Why (and how) we've switched away from google maps. Online at the Nestoria official blog: <http://blog.nestoria.co.uk/why-and-how-weve-switched-away-from-google-ma>.
- OpenStreetMap (2012a). Osm: Map features: A summary of commonly used tags for main elements used to describe features within openstreetmap. The OpenStreetMap Wiki pages: http://wiki.openstreetmap.org/wiki/Map_Features.
- OpenStreetMap (2012b). What are openstreetmap mapping parties? Online at the OpenStreetMap Wiki: http://wiki.openstreetmap.org/wiki/Mapping_parties.
- OSM-Talk-List (2011). Mailing list conversation regarding a dispute caused by the bulk import of vector data in quebec, canada. OSM Talk Mailing Lists: <http://www.mail-archive.com/talk@openstreetmap.org/msg36440.html>.
- Over, M., Schilling, A., Neubauer, S., and Zipf, A. (2010). Generating web-based 3d city models from openstreetmap: The current situation in germany. *Computers, Environment and Urban Systems*, 34(6):496–507.

- Perkins, C. and Dodge, M. (2008). The potential of user-generated cartography: a case study of the OpenStreetMap project and mapchester mapping party. *North West Geography*, 8(1):1932.
- Poore, B. and Wolf, E. (2012). Metadata squared: Enhancing its usability for volunteered geographic information and the geoweb. In Sui, D., Elwood, S., and Goodchild, M., editors, *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, pages 43–64. Springer, Heidelberg, Germany.
- Prasarnphanich, P. and Wagner, C. (2011). Explaining the sustainability of digital ecosystems based on the wiki model through critical-mass theory. *Industrial Electronics, IEEE Transactions on*, 58(6):2065–2072.
- Ramm, F., Topf, J., and Chilton, S. (2010). *OpenStreetMap: Using and Enhancing the Free Map of the World*. UIT Cambridge, Cambridge, CB4 1GQ, England.
- Reagle, J. (2011). The argument engine. In Lovink, G. and Tkacz, N., editors, *Critical Point of View: A Wikipedia Reader. INC Reader 7*, pages 14–33. Institute of Network Cultures, Amsterdam, The Netherlands.
- Richter, C., Miscione, G., and Georgiadou, Y. (2010). Conceptualizing people in SDI literature: Implications for SDI research and development. *International Journal of Spatial Data Infrastructures Research*, 5:286–325.
- Richter, K.-F. and Winter, S. (2011). Citizens as database: Conscious ubiquity in data collection. In Pfoser, D., Tao, Y., Mouratidis, K., Nascimento, M., Mokbel, M., Shekhar, S., and Huang, Y., editors, *Advances in Spatial and Temporal Databases*, volume 6849 of *Lecture Notes in Computer Science*, pages 445–448. Springer Berlin / Heidelberg. 10.1007/978-3-642-22922-0_27.
- Smart, P. D., Quinn, J. A., and Jones, C. B. (2011). City model enrichment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(2):223–234.
- Suh, B., Convertino, G., Chi, E. H., and Pirolli, P. (2009). The singularity is not near: slowing growth of wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration, WikiSym '09*, pages 8:1–8:10, New York, NY, USA. ACM.
- Topf, J. (2012). Openstreetmap for land and water polygons and global coastlines. Online: <http://openstreetmapdata.com/>.

- Wang, L. S., Chen, J., Ren, Y., and Riedl, J. (2012). Searching for the goldilocks zone: trade-offs in managing online volunteer groups. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, pages 989–998, New York, NY, USA. ACM.
- Wen, Y., Chen, M., Lu, G., Lin, H., He, L., and Yue, S. (2012). Prototyping an open environment for sharing geographical analysis models on cloud computing platform. *International Journal of Digital Earth*, 0(0):1–27.
- Wikimedia (2012). Wikimedia foundation financial reports. Online at (http://wikimediafoundation.org/wiki/Financial_reports).
- WikipediaStats (2012). Statistics of contributions to wikipedia. Online at: <http://stats.wikimedia.org/EN/TablesWikipediaEN.htm>.
- Zielstra, D. and Hochmair, H. H. (2011a). Comparative study of pedestrian accessibility to transit stations using free and proprietary network data. *Transportation Research Record*, (2217):145–152.
- Zielstra, D. and Hochmair, H. H. (2011b). Digital street data: Free versus proprietary. *GIM International Journal*, 25(7):7–14.
- Zook, M., Graham, M., Shelton, T., and Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the haitian earthquake. *World Medical & Health Policy*, 2(2):Article 2.

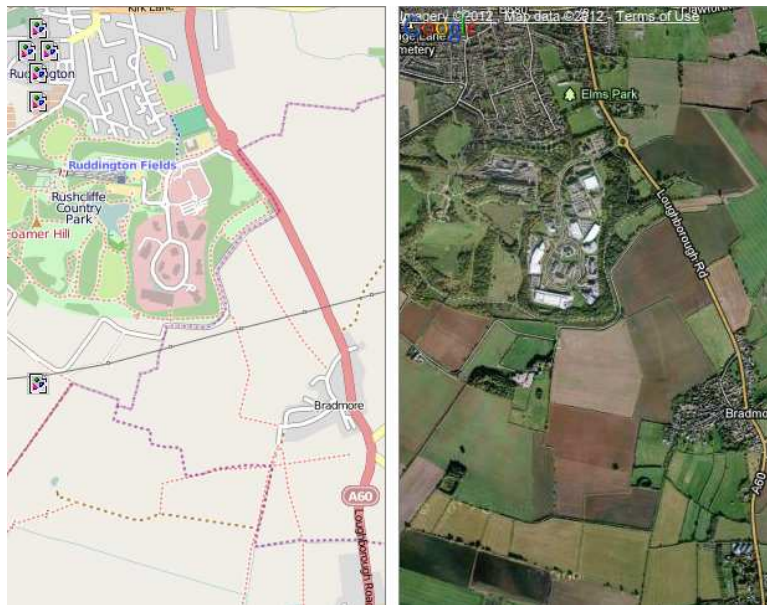


Figure 1: An example of urban meets rural in OSM in Nottingham, UK. The OSM map is compared with Google Aerial Imagery of the same location.

Characteristic	OpenStreetMap	Wikipedia
<i>Verifiability</i>	a tag/value combination is verifiable if and only if independent users when observing the same feature would make the same observation every time. Testable definitions are documented on the OSM Wiki	“The threshold for inclusion in Wikipedia is verifiability, not truth, that is, whether readers are able to check that material added to Wikipedia has already been published by a reliable source, not whether we think it is true” (Reagle, 2011)
<i>Administrators</i>	Not known however small in number	English Wikipedia has approximately 1,500 administrators
<i>Sustainability</i>	Good sustainability: Recent change of OSM license structure (May 2012) caused friction amongst many “senior” mappers (Neis and Zipf, 2012). Slowly gaining some influential industry support. Must address issues of quality and validation before more mainstream acceptability.	High sustainability: Prasarnphanich and Wagner (2011) uses empirical evidence to show that ‘Core’ contributors (high frequency contributors) are transient but new ‘Core’ groups constantly emerge from more peripheral participation levels. The “cost” of adding new information must remain low (Greenstein, 2012)
<i>Admin Status</i>	Members assume responsibility for an spatial area or issue - must be trusted by the community and seen as a productive contributor	No formal process exists as one can assume these positions. There is a community review process to evaluate suitability.
<i>Dispute Detection</i>	Detected as rapid changes to values of tags or feature geometries and are considered vandalism. Community vigilance and bot agents assist early detection.	Most notably the “three-revert-rule” where an editor cannot perform more than 3 reverts on a single page in 24 hrs. Community vigilance and bot agents assist early detection.
<i>Dispute Resolution</i>	The “On-the-ground” rule is used giving precedence to the names/attribute values people use locally in the disputed area. OSMF Data Working Group can help with arbitration, if necessary.	Focus must remain on the page content and not personally against other editors. Discussion, mediation, third-party intervention, etc. are usually successful in solving most disputes. An arbitration committee is elected each year.
<i>Contents</i>	141 million ways	Approximately 4 million articles
<i>Featured Objects</i>	6 million (about 4%) with greater than 15 edits (Mooney and Corcoran, 2012)	Currently, 3,586 featured articles

Table 2: A comparison of some of the collaborative editing strategies of Wikipedia and OpenStreetMap. Statistical data are correct as of November 2012