Placing Wikimapia: An exploratory analysis

Andrea Ballatore
Department of Geography
Birkbeck, University of London, UK
a.ballatore@bbk.ac.uk

Jamal Jokar Arsanjani
Department of Development & Planning
Aalborg University, Copenhagen, Denmark
jja@plan.aau.dk

International Journal of Geographical Information Science (2018) – Author copy

Abstract

Wikimapia is a major privately-owned volunteered geographic information (VGI) project to collect information about places. Over the past ten years, Wikimapia has attracted hundreds of thousands of contributors and collected millions of data points, including towns, restaurants, lakes, and tourist attractions (http://wikimapia.org). Unlike OpenStreetMap, Wikimapia adopts a "placial" perspective, favouring rich descriptions over detailed geometries and encouraging the collection of textual and visual content about places with approximate footprints. In this article, we first trace the origin and development of Wikimapia as a for-profit project, intimately linked with search engine advertising. Drawing on an in-depth interview with a former developer, we analyse project's data model and characteristics of its community. As Wikimapia discussions are rife with copyright issues, we discuss the project's intellectual property, as well as its strategies for quality management. Second, we focus on the popularity of the project, which is crucial to the longevity and sustainability of VGI projects. Using behavioural data from Google Trends, we trace a geography of interest in Wikimapia, comparing with that in OpenStreetMap, from a temporal and spatial perspective. While OpenStreetMap attracts more interest in high-income countries, Wikimapia emerges as relatively more popular in low- and middle-income countries, countering the received notion of VGI as a Global North phenomenon. Our study suggests that Wikimapia's popularity is steadily declining.

Keywords: Wikimapia, OpenStreetMap, volunteered geographic information, user-generated content, data quality, Internet geography, crowdmapping, crowdsourcing

1. Introduction

A decade after Goodchild's seminal article (2007), volunteered geographic information (VGI) has become a fundamental part of our informational ecosystem. Through their digital devices, citizens have become a prominent source of geospatial information, producing data about social and natural phenomena in active, passive, and, perhaps more importantly, unknowing ways. VGI-related trends have been discussed with many terms, including neogeography, crowdmapping, spatial user-generated content, and citizen science, capturing their different facets and complexities (See et al., 2016).

Every day, Google Maps routinely collects geographical information from its billion users, both silently and interactively; TripAdvisor and Airbnb crowdsource the opinions of millions of users about places; many mainstream online services use OpenStreetMap data as part of their base maps; after every natural disaster, open maps of vulnerable areas are rapidly assembled. As examples of this kind abound, it can be argued that VGI is losing its novelty. As the sociologist of technology Vincent Mosco (2004) pointed out, it is precisely when new, successful technologies "withdraw into the woodwork" of everyday banality that their effects become real and profound (p. 2). Therefore, as its power and reach grow, research into VGI is more important than ever.

Despite the throng of published research articles, fundamental questions about VGI remain only partially answered and persistently escape broad generalisations. Who are the contributors and why do they engage in spatial information production, and what incentives work or do not work (Budhathoki & Haythornthwaite, 2013)? How can we calculate the quality and fitness-for-purpose of crowdsourced data in a reliable, preferably intrinsic way (Goodchild and Li, 2012)? What are the limitations of such models and what are their spatial, epistemic, and cultural biases (Dodge & Kitchin, 2013)? How can we design peer production models so that they are more inclusive towards under-represented groups and regions (Stevens et al., 2014)? What aspects of spatio-social processes can VGI help understand, for example in transportation science, urban planning, public health, sociology, political science, and the humanities?

Another important question, we believe, is whether our inventory of VGI projects is sufficiently broad and deep to appropriately indicate what can, cannot, and should not be done with them. In this sense, it is not hard to see that academic attention has been unevenly distributed in the VGI landscape: While a remarkable amount of studies and applications have expounded OpenStreetMap, crowd-mapping project Wikimapia[1] is mentioned in passing as an important example of VGI (e.g., Goodchild, 2007; Elwood et al., 2012). A crude but effective indicator of academic interest is the number of results on Google Scholar, which currently (August 2017) returns 24,300 hits for OpenStreetMap and 3,700 for Wikimapia.[2] To date, few studies focussed explicitly on Wikimapia, pointing out how it manages cartographic vandalism (Ballatore, 2014), how it seems to favour Arab content over Jewish in Jerusalem, going against the general bias towards high-income regions found in much VGI (Bittner, 2016), and how it can be used to track toponymic changes in the Middle East (Ahmouda & Hochmair, 2017).

In this study, our first objective consists of analysing the origin, characteristics, scope, and limitations of Wikimapia as a VGI source, with a particular focus on data production and quality control. An extensive interview with a former developer of the project provided several insights that help frame Wikimapia data and policies in its context (Milevski, 2016).[3] Secondly, we target the online visibility of Wikimapia and we compare it with that of OpenStreetMap over time and space, with surprising results. The online visibility of these projects on search engines is a central element to explain their ascent or decline in different parts of the world and, to this

[1] http://wikimapia.org (accessed on 10 February 2018)
[2] Similar ratios can be obtained on other scholarly engines, such as Microsoft Academic and CiteSeer.
[3] For this study, we also contacted the project owners and other team members, who were unavailable.

purpose, we provide a comparative geography based on Google search data. All the datasets used in the study are available online.[4]

The remainder of this paper is organised as follows. Section 2 surveys the related work on VGI, framing this study on Wikimapia in ten years of scholarship. Section 3 provides an overview of Wikimapia origin, ethos, community, copyright, and quality management. An analysis of Wikimapia and OpenStreetMap online popularity, followed by a geography of online interest, is outlined in Section 4. Finally, Section 5 provides conclusive remarks and directions for future research.

2. Related work

After ten years of rapid development, phenomena described as VGI, user-generated content, spatial crowdsourcing, and crowd-mapping, have gained prominence in geographic information science (GIScience), and play a major role in supporting a wide range of applications, providing troves of geographic data with light licensing commitments. Among the existing platforms for collecting and sharing geographic data, OpenStreetMap has indeed proven itself to be a long-lived platform, with a growing number of contributors and data. It is fair to argue that OpenStreetMap served as a catalyst in GIScience, spurring new research directions for researchers, challenging traditional, closed, commercial, and authoritative data providers, and promoting new open data policies across the world (Jokar Arsanjani et al., 2015). This can be explained by its unique and dynamic mechanisms for data collection, transparent policies, clear licensing and copyright management, flexible tools for editing and retrieval of spatial information.

OpenStreetMap research

The research landscape on OpenStreetMap has been very extensive and a wide range of topics have been studied. Among others, there have been efforts on quality assessment (Neis et al., 2014; Glasze & Perkins 2015), user incentives and editing behaviour (Haklay et al., 2010; Stein et al., 2015), socio-economic aspects of contributions (Jokar Arsanjani & Bakillah 2015; Neis et al., 2013), disaster management through humanitarian mapping (Palen et al., 2015), and land management (Kalantari et al., 2015). Major efforts have looked into data quality since it is paramount to use VGI data in real applications (Goodchild & Li, 2012), such as intrinsic (Barron et al., 2013) and extrinsic quality measures (Jokar Arsanjani & Fonte 2016). Novel approaches include the notion of a contribution index (Jokar Arsanjani et al., 2016) and of conceptual quality (Ballatore & Zipf, 2015). A recent survey by Senaratne et al. (2017) provides an extensive review of quality assessment techniques for image-based, map-based and text-based VGI.

As argued by Muki Haklay,[5] the excessive identification between VGI and OpenStreetMap is problematic and even stifling for GIScience. A more ecological view of the different VGI projects

[4] https://github.com/andrea-ballatore/WikimapiaResearch (accessed on 6 April 2018)
[5] https://povesham.wordpress.com/2014/08/14/openstreetmap-studies-and-why-vgi-not-equal-osm (accessed on 30 Aug 2017)

as an inter-connected network is needed. In the same decade, similar projects emerged, providing alternative platforms for collecting geospatial data, such as Wikimapia, Wikidata, and Platial (now closed). Each project is built on a different ecosystem, adopting data models, policies, licenses, with its own community of contributors and users with different objectives and interests. Therefore, studying platforms beyond OpenStreetMap and identifying their importance in GIScience is a crucial task. As Wikimapia is one of the oldest and still existing projects, an in-depth analysis of the project is long overdue.

A divided VGI world

VGI is produced through social processes and therefore its social dimension and dynamics should always be in the spotlight, drawing on research in human geography. Beyond the obvious contribution inequality that characterises peer-production systems (i.e., few very active users produce most of the content), digital divides play an opaque role in shaping VGI. Unequal patterns of access to digital networks, of participation to projects and platforms, and of representation of different groups result in specific biases, detectable in the datasets (Graham et al., 2015b). Hence, any discussion on data quality and fitness-for-purpose must take into account the thematic, demographic, and geographic biases that are inevitably reflected in the datasets. Sadly, as pointed out by Elwood et al. (2012), "underlying structural inequalities remain unaltered" (p. 583) and there is little reason to think that things will depart from the current trajectory.

In their extensive survey, Graham et al. (2015b) documented the informational dominance of the Global North in sources as diverse as Wikipedia, GitHub, Google searches, OpenStreetMap, and Freebase. More specifically, Graham & De Sabbata (2015) showed how open gazetteer GeoNames has an uneven density of named places, skewed towards countries that tend to be wealthy and populous, and that support open data policies. Although not strictly a mapping project, Wikipedia is similarly produced in a regime of uneven participation, with five high-income countries generating almost half of all edits (Graham et al., 2015a). In light of these considerations, the fact that Wikimapia does not seem to follow this Global-Northern bias in Jerusalem (Bittner, 2016) provides further reasons to further investigate this project.

3. Wikimapia: An overview

The Wikimapia project was founded by Russian entrepreneurs Alexandre Koriakine and Evgeniy Saveliev in Moscow in 2006. Earlier that year, Flickr introduced a service where it was possible to tag an area of a picture. According to a former developer, [6] one of the founders had the idea that this could be extended to maps, to tag places on satellite imagery (Milevski, 2016). Initially, the project had no external funding and started to get traction on online boards (Milevski, 2016). The core idea consisted in asking users to draw rectangles on a satellite background, and add descriptions of places, such as cities, parks, and notable buildings, adopting the Wikipedia

[6] Alexander Milevski gave us explicit authorization to use his name in this article at the beginning of the interview (Milevski, 2016).

approach in a geospatial dimension. While the project never used the term, its core informational asset is a crowdsourced digital gazetteer, unlike the complex vector data created by OpenStreetMap. The project's goal was, and still is, "to describe the whole world by compiling as much useful information about all geographical objects as possible."[7]

Data model

The core entity in the Wikimapia ecosystem is the *Place*, also called *Object*. Over its 11 years of activity, Wikimapia collected data about millions of geographical objects, of which 2.3M houses, 1.2M villages, 1.2M shops, and 0.9M schools.[8] The total number of these objects rose from 10M in 2009 to 23M in 2015, and is currently about 27M.[9] Initially, users were asked to characterise each Place with a set of categories and a bounding box, outlined from Google satellite imagery. From a conceptual perspective, similarly to OpenStreetMap, Wikimapia adopts an object view of the spatial entities it wants to model (Kuhn, 2012). Subsequently, bounding boxes became polygons, which can be drawn on the map editor. The creation of polygons is considered to be highly entertaining by users (Milevski, 2016). It is interesting to note that points, a simple and intuitive representation of places, are not included at all in the data model. Figure 1 shows a typical Wikimapia object (Gorky Park in Moscow).

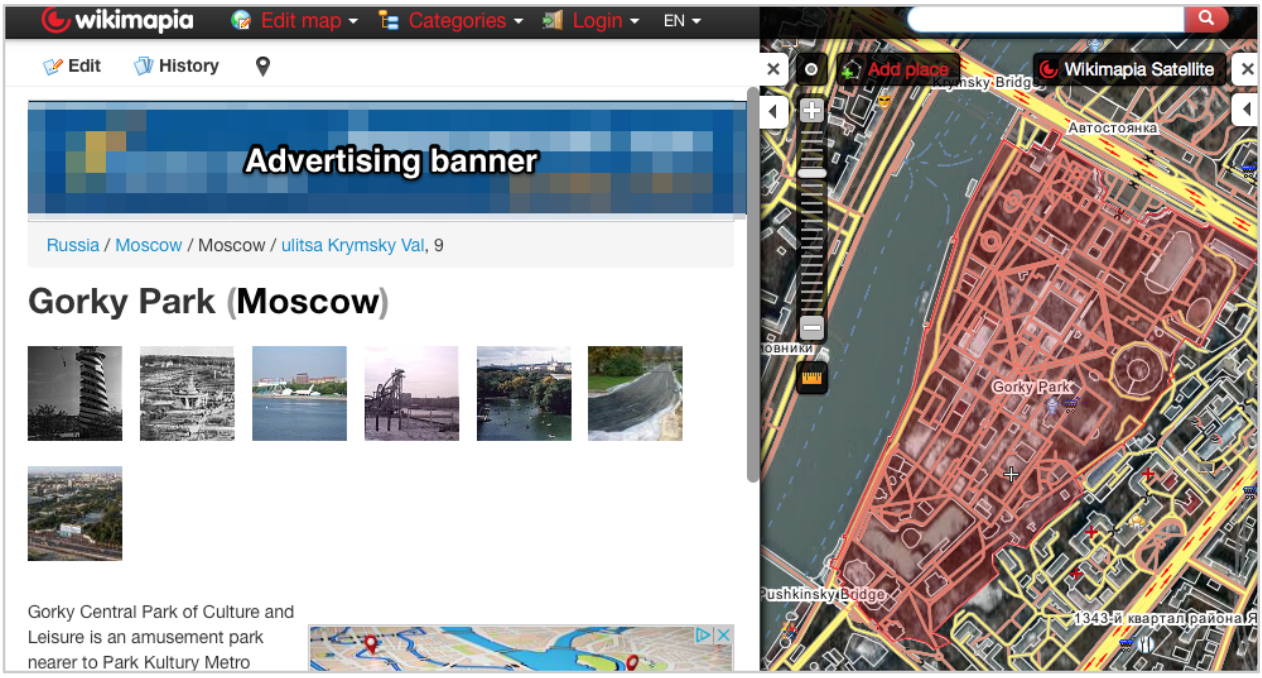

Figure 1: The Wikimapia object for Gorky Park in Moscow, available at http://wikimapia.org/348/Gorky-Park (accessed on 25 August 2017)

OpenStreetMap adopts a loosely regulated set of key-value pairs called tags to describe geographic objects (Ballatore & Mooney, 2015). By contrast, Wikimapia's semantic approach

[7] http://wikimapia.org/about (accessed on 24 July 2017)
[8] http://wikimapia.org/#show=/stats/ (accessed on 17 June 2017)
[9] http://wikimapia.org/#show=/stats/ (accessed on 24 July 2017)

hinges on a centrally-curated taxonomy of categories.[10] Examples of categories include *park -> garden -> botanical garden* and *railway -> train station -> railway terminal*. As largely expected, a small set of categories, including *house*, *building*, and *village*, describe most places in the dataset. The creation and deletion of categories are highly regulated, and only users who have gained administrative rights can add or remove new categories.[11] When a category is semantically overlapping with others, and its usage is limited, it is flagged as "unapproved". Following Wikipedia, Wikimapia asks to adopt a neutral point of view (NPOV) in the geographic information being created. The data is currently only available through an API, which does not allow full data dumps. There are no full data dumps because, in the early 2010s, the entire website was maliciously scraped to set up a parallel site, earning revenue from advertising (Milevski, 2016).

In addition to categories, places have other attributes that include photographs and comments (see for example the Iranian capital Tehran).[12] Although Wikimapia is more placial than OpenStreetMap, the project has included linear features, such as roads, rivers, and railways, represented as categorised polylines. However, these kinds of objects did not enjoy a growth comparable to places, and are still considered a beta feature of the platform. The only other data source that is interlinked with Wikimapia places is Wikipedia. For copyright reasons, OpenStreetMap and Wikimapia data are explicitly kept separate, and no imports from one to the other are either allowed or encouraged.[13]

Community

While it is very hard to estimate the number of active users, the total number of registered users rose from 400,000 in 2009 to over 2.5M in 2017. The number of new registered users has been steadily declining from more than 400,000 in 2010 and in 2011 to 110,000 in 2016, still amounting to a significant group of contributors. Never more than 7 people were professionally involved in developing the project in the company. Currently, all team members are Russian.[14]

Contributors, according to the Wikimapia developer (Milevski, 2016), either tag primarily their local areas, or explore other areas for travelling purposes. The platform is used by some photographers and travellers to record travel information, even to organise travelogues. In the developer's view, contributions originated primarily from India, Russia, United States, and Latin America, while the project has not become as popular in Western Europe, although some top users are German and English. The data is used by small business predominantly in the Middle East, India, and in the US to promote their online visibility, as well as their spatial and thematic findability.

[10] http://wikimapia.org/user/tools/categories_catalog (accessed on 24 July 2017)
[11] http://wikimapia.org/wiki/User_Guide:_Places_and_Place_Tags (accessed on 24 July 2017)
[12] http://wikimapia.org/5622633/Tehran (accessed on 17 June 2017)
[13] https://wiki.openstreetmap.org/wiki/Wikimapia (accessed on 24 July 2017)
[14] http://wikimapia.org/team (accessed on 15 August 2017)

Milevski stated that, in demographic terms, the average age of contributors is around 37 years, and the gender is predominantly male (Milevski, 2016). As expected in peer-production systems, the top contributors in the community have a considerable weight on the overall volume of contributions. By directly observing the accounts of the top 20 Wikimapians,[15] it is possible to note that they are predominantly male (19 out of 20), and tend to be English, German, Russian, and Ukrainian-speaking. From 2008, the top user has made 1.4M revisions to the dataset. More empirical investigations are needed to quantify the geographic and national biases in the placial data production process.

The contributors form a complex society, with a hierarchical structure, power dynamics, and sub-groups. The developer stated that Wikimapians "had revolutions and plots", for example with groups trying to block and ban other groups of users, when disagreeing on editing choices (Milevski, 2016). In response to these social turbulences, the company had to create committees to draw formal policies and investigate severe incidents. As described more in detail below, and unlike similar projects, the company created a complex system to manage and encourage users with experience points, awards, and privileges, in the general paradigm of gamification.

Copyright and business model

Wikimapia is radically different to more well-known projects like OpenStreetMap and Wikipedia in a number of ways. Being for-profit, Wikimapia has since its inception experimented business models to monetise the visits to the website. Advertising has been the main source of revenue for the company. Wikimapia places are associated with permalinks, and each page is visualised with contextual ads from Google AdSense. For example, when visualising the entry for Moscow, the page shows a Google ad about travel offers to reach the city, as well as a generic ad of a miracle diet.[16] In the early 2010s, the company also developed its own spatial advertising model, aiming at small businesses.[17] However, Milevski (2016) stated that this approach never took off because, while potential clients liked the idea of appearing more prominently on the map, were not willing to pay for it.

The project faces copyright-related tensions from two perspectives, both mingled with its relationship with Google. First, its main data source was Google satellite imagery, mainly used to identify and generated the footprints of spatial objects. Wikimapia data, therefore, can be classified as "derived work" in many jurisdictions, even though facts about phenomena on the surface of the Earth cannot be copyrighted. The lack of an explicit agreement with Google cast a shadow on the legal standing of the company's products, which remains unclear to date.[18]

[15] http://wikimapia.org/#m=b&show=/user/tools/users_rating/ (accessed on 25 July 2017)
[16] http://wikimapia.org/5802505/Moscow (accessed on 24 July 2017). The ads are personalized, so different ones might appear to different users.
[17] http://wikimapia.org/ads (accessed on 24 July 2017)
[18] https://wiki.openstreetmap.org/wiki/Wikimapia (accessed on 24 July 2017)

Second, the ad-driven business model caused significant reputational problems to the company among open-source and peer-production activists, who accused it of exploiting volunteered labour for private profit. This is a recurring issue in peer-production platforms, where digital labour is often unrewarded (Graham & Anwar, 2018). To settle some of the controversies, in 2012 the company changed the license of their core dataset to Creative Commons license Attribution-ShareAlike (CC BY-SA), adding "provide free access to our data for public domain" to their mission.[19] However, the legal ground of the license with respect to the Google-derived data remains controversial. Hence, Wikimapia data cannot be fully considered part of the open data, commons-based ecosystem that includes Wikipedia, OpenStreetMap, and GeoNames. This, as discussed below, does not imply that the project's data and web pages do not create value for its users, particularly in developing countries.

Quality management

As Wikimapia aims at ad revenue, data quality is intrinsically linked with the visibility of the platform. Hence, the project developers managed the platform to attract meaningful contributions through an explicit user management strategy. They focused on facilitating the place creation and editing process, rather than aiming at a detailed, formal data quality criteria that, Milevski (2016) claimed, will arguably never be possible through a peer production model. As the developer stated, the most active contributors tend to want to restrict everything to ensure high quality, but such an approach would increase barriers to entry for new contributors, so the project management decided to keep the model more open, while enabling some restrictions for novice users, in a typical trade-off found in peer production systems. In his discussion of crowdsourcing, Brabham (2013) pointed out that fun, connectedness, and peer feedback are the three central motivational mechanisms to keep users involved, and the Wikimapia team adopted them all in designing the platform.

At the core, Wikimapia hinges on a reward mechanism for its users, with a stronger user hierarchy than in OpenStreetMap. In the platform design, major importance was given to gamification, to user rankings, and to an award system with more than 50 awards for high-quality contributions. By creating or editing places, users gain *experience points*, a concept deriving from role-playing games. Users have a *level*, starting at 0.[20] Positive contributions are translated into experience points, and the user level is determined automatically based on their experience points. Different actions translate into different experience points, e.g., creating a polygon is worth 50 points, while adding an address is worth 15. User levels range from 1 to 8, with negative values for problematic and banned users. For instance, to reach level 3, 15,000 experience points are necessary.

Users can vote for each other, publicly endorsing valuable contributions. As a result of consistent, outstanding contributions, users can be promoted to the status of *advanced users*, who gives access to moderating tools and actions on the platform, such as protecting places from editing. The promotion is carried out through co-optation by other advanced users on the

[19] http://wikimapia.org/forum/viewtopic.php?f=74&t=9878 (accessed on 24 July 2017)
[20] http://wikimapia.org/wiki/User_Guide:_User_levels (accessed on 27 July 2017)

user forum. To date, the board of advanced users include 534 people, and deals with several aspects of quality control, including vandalism management (Ballatore, 2014). Gamification is not without its critics, and several users actively opposed the idea of user ranking, suggesting it was introducing competition in a leisure context.[21]

To keep users engaged, the platform uses gamification also by distributing a range of awards that appear on user profiles.[22] Regular awards focus on content creation from a quantitative perspective, signalling for example that a user created more than 100 places, while special awards are thematic and quirky. For example, among 53 special awards, the badge *Hedonist* is awarded after adding 20 places in the dining category, and *Mad Max* is given to users that map vehicle-related places (see Figure 2).

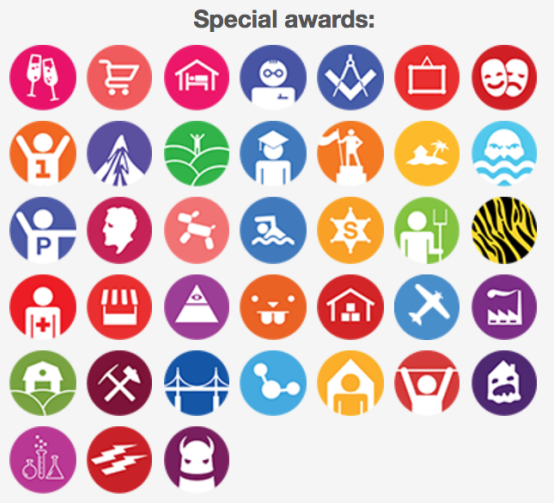

Figure 2: Wikimapia special awards (accessed on 27 July 2017)

While spatial accuracy is not given priority, a central concern for data quality in Wikimapia is completeness, trying to reduce the unevenness of the spatial distribution of places. A central tool that was created by the Wikimapia team in 2011 for this purpose is the Status Grid.[23] As shown in Figure 3, this tool shows statistics for each cell, including the number of places without categories and photos. The grid cell size varies at different scales. A formula was devised by the team and by the community to summarise the aspects into a score from 0% (low completeness) to 100% (high completeness).[24] The formula is a weighted linear combination, based on intrinsic metrics, similar to many metrics proposed for OpenStreetMap (e.g., Senaratne et al., 2017). The tool obtained mixed results in terms of fostering completeness, also due to software development issues that seemed to hamper its usability.

[21] http://wikimapia.org/#lang=en&lat=-58.175048&lon=-156.288757&z=9&m=b&show=/11453897/Rank-Awards-Dump (accessed on 12 February 2018)
[22] http://wikimapia.org/user/tools/awards (accessed on 27 July 2017)
[23] http://wikimapia.org/docs/Status_grid (accessed on 27 July 2017)
[24] http://wikimapia.org/forum/viewtopic.php?f=5&t=8900 (accessed on 27 July 2017)

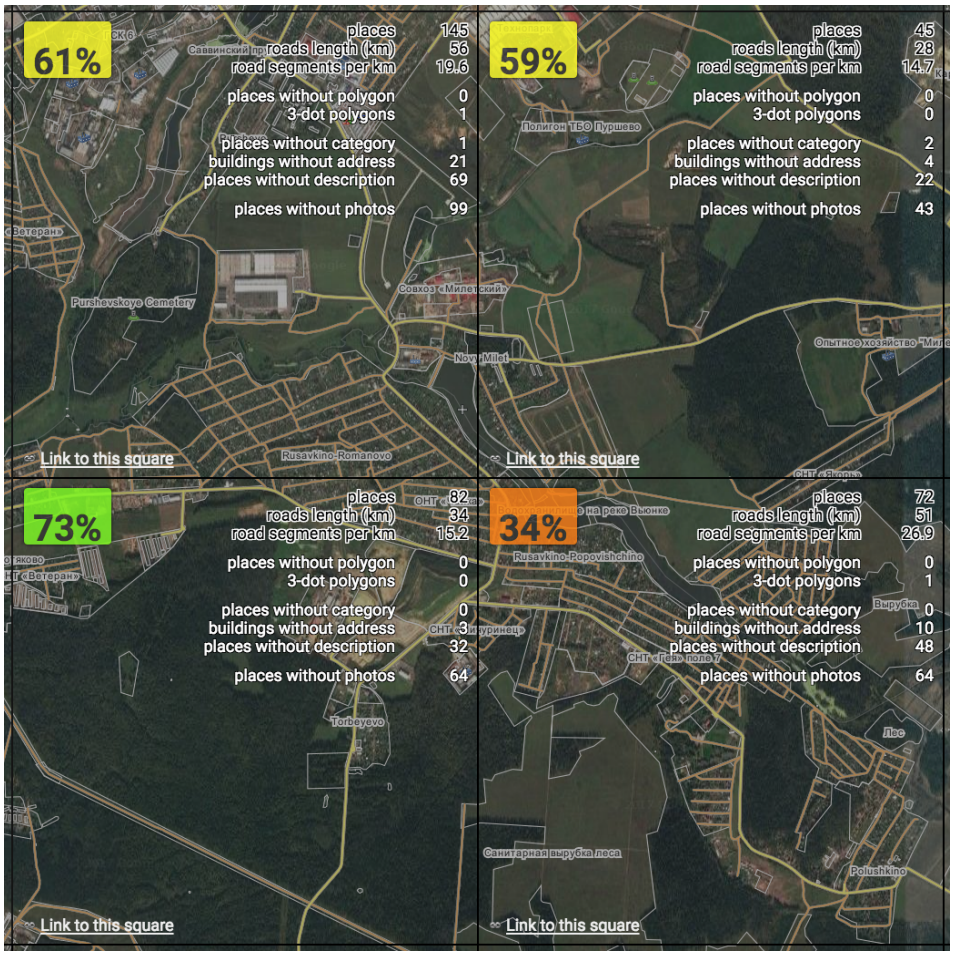

Figure 3: The Grid Status tool, visualising completeness statistics for four cells located west of Moscow (accessed on 27 July 2017)

4. Web visibility and search geography

In line with their business model, the Wikimapia team actively promoted the visibility of the platform on the web, particularly through search engines. Search engine optimisation (SEO) was deployed early to produce pages representing places that would rank highly in Google search results. Google users searching for place names would view Wikimapia pages, particularly when the places are not well mapped on Google Maps or competing, mainstream products. While precise data about views is not available, at its peak in 2009–2010, each day the project boasted about 3M page views, and about 1M unique visitors, sustaining the project financially.[25] In the remainder of this section, we conduct an exploratory analysis of the Web popularity of Wikimapia, comparing it with OpenStreetMap.

[25] http://wikimapia.org/ads (accessed on 24 July 2017)

Web popularity of Wikimapia and OpenStreetMap

To study the popularity of these projects online, three orthogonal aspects can be observed: (1) the number of visits to their websites, (2) their ranking in search engines, and (3) the volume of searches for their names. For the first aspect, it is useful to observe Wikimapia and OpenStreetMap's Alexa rankings. Alexa Internet, Inc. produces one of the most popular indicators of the popularity of websites based on estimates of visits, commonly used in Web science (e.g., Janc, 2016). At the time of writing (December 2017), Wikimapia[26] is ranked 4,993 globally, and 773 in Russia (the country in which it is most popular), while OpenStreetMap[27] is ranked 6,487 globally, and 1,508 in Germany, where the project attracts its largest community. To put this data into perspective, Wikipedia ranks 5th globally, and MapQuest 2,030.[28] Interestingly, despite OpenStreetMap having more contributors and more complex data, Wikimapia still attracts more visits, probably through its higher visibility on Google search results. However, according to Alexa data, Wikimapia's visibility is experiencing a steady downward trend from about 3,000 to the current 4,993, confirming the general decline of the project. On the other hand, OpenStreetMap displays a more unstable pattern, and is currently on the rise, suggesting that the project might become more visible in the next few years.

While the ranking of websites is not directly observable, Google searches provide a proxy to estimate the online visibility of these projects. While it is hard to estimate the overall visibility of a website on the search engines, Google Trends[29] provide aggregate "big data" about user search behaviour. The service does not show absolute volumes of searches, but scales the data to allow for comparisons between searches, showing the relative interest in topics in the total pool of searches over time and geographic space. It is important to bear in mind that this data refers to search behaviour, and not to actual data production and consumption on Wikimapia and OpenStreetMap. Additionally, Internet access and Google search engine use vary nationally, reaching widely different percentages of the population (Ballatore et al., 2017). While acknowledging the limitations and noise contained in the data, Google Trends has been successfully used to investigate economic and cultural behaviours at a large scale (Preis et al. 2012; Choi & Varian, 2012; Stephens-Davidowitz, 2013).

From a methodological perspective, the selection of queries is crucial to extract meaningful signals. To ensure replicability, we specify the URLs through which the data was collected. Google Trends collects data about individual, case-insensitive strings or topics, which are clusters of semantically related strings. Strings and topics are not directly comparable, as they correspond to data aggregated using very different approaches. A topic exists for OpenStreetMap,[30] but not for Wikimapia, limiting the analysis to strings. The most popular spellings for the projects, "Wikimapia" and "OpenStreetMap", were therefore selected as input strings.

---

[26] http://www.alexa.com/siteinfo/wikimapia.org (accessed on 27 July 2017)
[27] http://www.alexa.com/siteinfo/openstreetmap.org (accessed on 27 July 2017)
[28] https://www.alexa.com/siteinfo/wikipedia.org (accessed on 30 August 2017)
[29] http://trends.google.com (accessed on 29 July 2017)
[30] https://trends.google.com/trends/explore?q=%2Fm%2F08f_0k (accessed on 6 December 2017)

By observing the popularity of Wikimapia and OpenStreetMap from 2004 to 2017, Google Trends shows a granular picture (see Figure 4).[31] Much search activity is strongly seasonal (Choi & Varian, 2012), and both trend lines exhibit remarkable seasonality, with major peaks in summer months and troughs in winter. This suggests a strong correlation between summer outdoor activities and holidays and geographic information search – at least in the Northern hemisphere. Small increases are also observable in November-December, suggesting either geographic search behaviour for Winter holidays in the Northern Hemisphere, or summer holidays in the Southern Hemisphere.

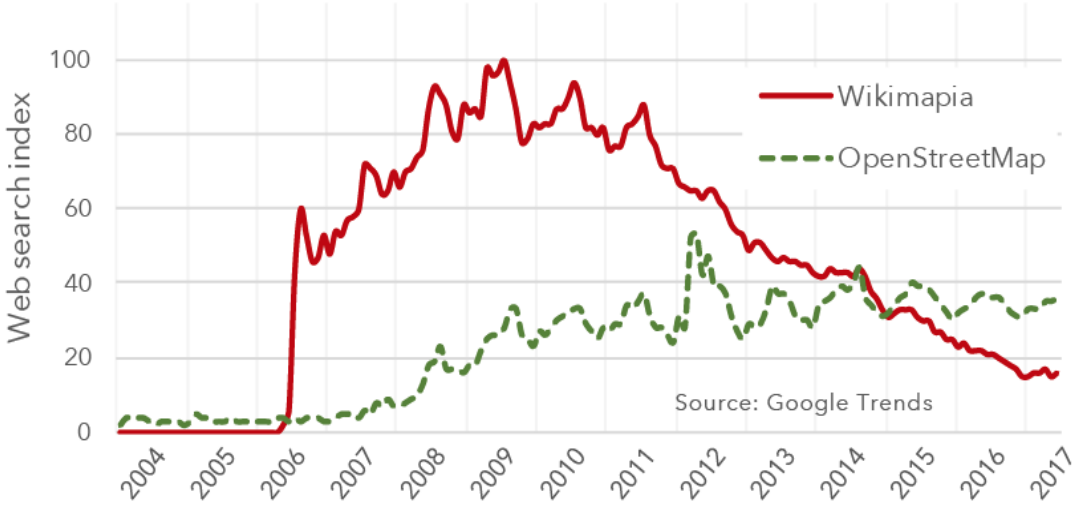

Figure 4: Google searches for Wikimapia and OpenStreetMap on a monthly basis (source: Google Trends worldwide from 2004 to 2017, accessed on 15 April 2017).

In terms of overall volume of searches (Figure 4), Wikimapia experienced rapid growth from 2006 and 2009, and then started a stable decline from 2012. By contrast, OpenStreetMap grew very slowly from 2004 to date. While Wikimapia was 5 to 6 times more visible than OpenStreetMap from 2006 to 2011, OpenStreetMap overtook it in 2014. This suggests that Wikimapia lost its upward trajectory in 2012, either because of lack of investment in SEO for the platform, or because of changes in the online geographic information markets. Although these trends are very clear, it is important to note that this data refers to Google searches, and not to direct usage of the websites.

To contextualise the search volumes for these user-generated projects, it must be noted that Google Maps shows a search volume that is at least 1,000 times higher than both Wikimapia and OpenStreetMap throughout their histories. Notably, Google Maps has reached 1B users

[31] https://trends.google.com/trends/explore?date=all&q=Wikimapia,OpenStreetMap (accessed on 20 May 2017)

worldwide, and its mobile app is the most used smartphone application.[32] So, while Wikimapia and OpenStreetMap are impressive and even dominant projects in the realm of user-generated spatial content and peer-production systems, they are definitely not able to compete with major corporate players that produce the bulk of mass interaction with spatial information online.

A geography of Wikimapia and OpenStreetMap searches

Being about crowdsourced geographic information, the completeness and biases of Wikimapia data are linked to where its contributors are located. Hence, we observe the spatial distribution of its searches around the world, comparing it with OpenStreetMap, bearing in mind the methodological limitations discussed above. It is also important to note that searches do not necessarily result in contributions, and that Google searches are only one entry point to these projects. Using Google Trends aggregate data as a proxy for interest in projects, it is possible to observe the relative volume of searches for Wikimapia and OpenStreetMap in each country and city, scaled between 0 (no searches) to 100 (highest searches).

In this data, a higher score means a higher proportion of all queries, not a higher absolute query count. This weighs the prominence of searches based on the total volume of searches, therefore avoiding large countries and cities to be over-represented.[33] Table 1[34] shows the Google Trends index scores from 2004 to date at the country level, for the top 5 countries for the two projects. As this indicates a radically divergent geography, with Wikimapia being popular in Arab-speaking countries and OpenStreetMap in German-speaking ones, it is worth observing this phenomenon globally.

Top countries	Wikimapia web search index	Top countries	OSM web search index
Morocco	100	Germany	100
Qatar	94	Austria	75
United Arab Emirates	73	Switzerland	46
Philippines	66	Belgium	44
Bahrain	56	Czech Republic	41

Table 1: Top five countries based on Google Trends web search from 2004 to 2017 (accessed on 15 April 2017).

[32] https://www.theverge.com/2017/5/17/15654454/android-reaches-2-billion-monthly-active-users (accessed on 29 July 2017)
[33] The user guide states: "A higher value means a higher proportion of all queries, not a higher absolute query count." For example, a small country where 2% of the queries are "Wikimapia" will get twice the score of a large country where "Wikimapia" obtains only 1% of the queries.
[34] Data based on URL https://trends.google.com/trends/explore?date=all&q=Wikimapia,OpenStreetMap in the "Region of interest" section (accessed on 17 June 2017).

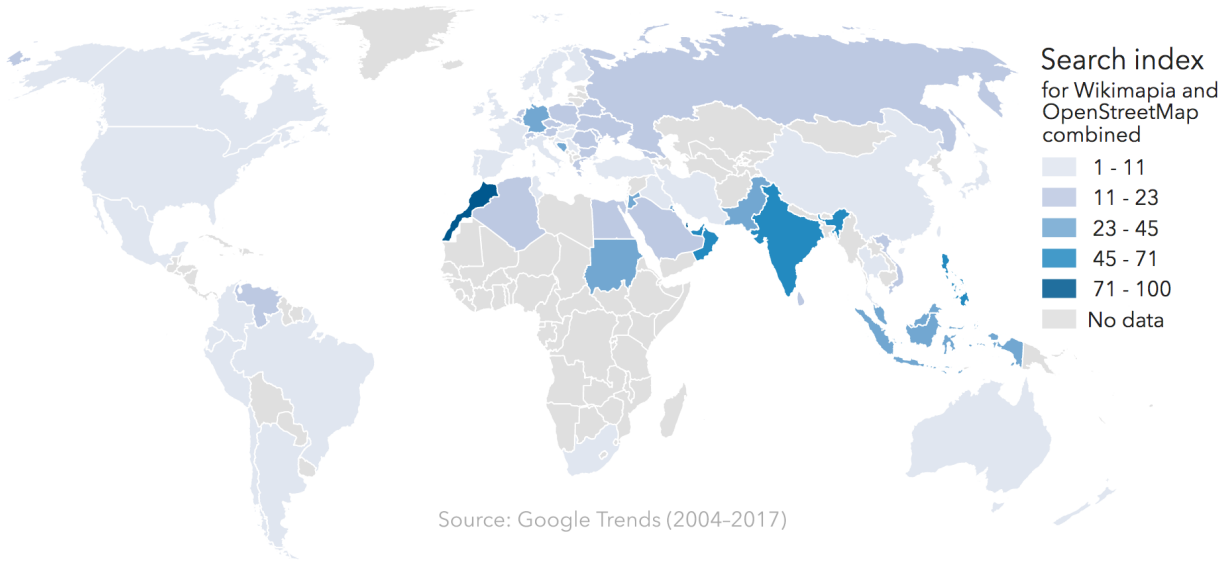

Figure 5: Combined search indices for Wikimapia and OpenStreetMap, binned with natural breaks (Jenks). This represents the overall relative interest in the projects in a country, with the respect to all searches. The top group (71-100) includes Morocco and Qatar, while the countries where the two projects have the lowest search index are China, Thailand, Peru, and Japan. Country boundaries from naturalearthdata.com, 2017.

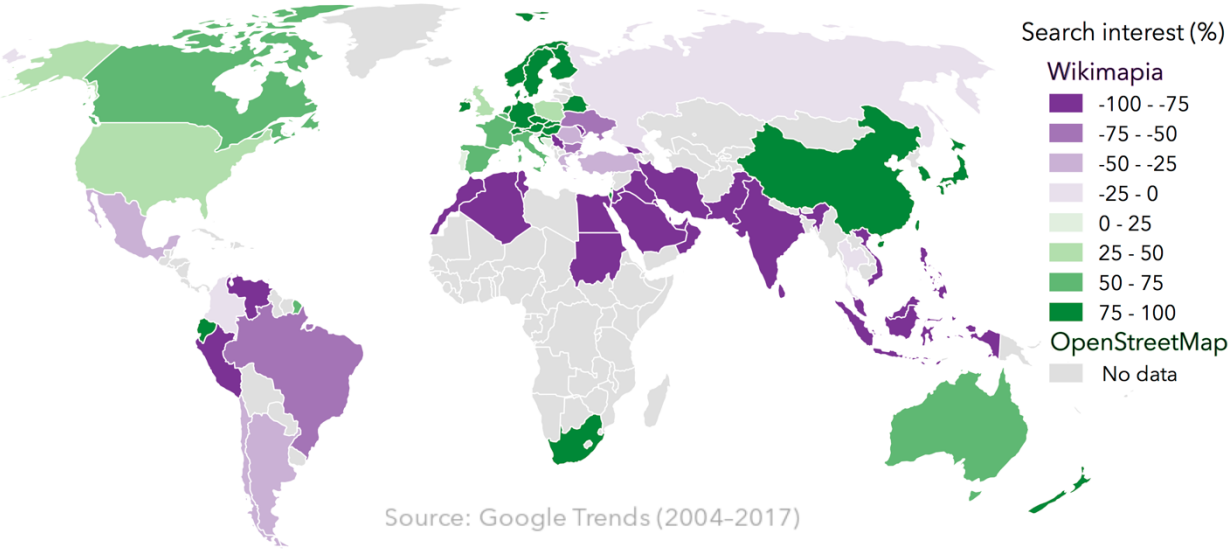

Figure 6: Search index for Wikimapia (purple) or OpenStreetMap (green), based on Google Trends from 2004 to 2017, as eight equidistant bins (accessed on 15 April 2017). The value is a percentage, ranging from -100% (all searches just for Wikimapia) to 100% (all searches just for OpenStreetMap). 0% means an even split between the two projects. Country boundaries from naturalearthdata.com, 2017.

Different countries exhibit a wide range of search indices, indicating varying levels of relative search interest in both Wikimapia and OpenStreetMap combined. The map in Figure 5 shows this variation, highlighting low levels of search in much of Western Europe, the Americas, China,

and Australia (index < 12). By contrast, Morocco, Qatar, United Arab Emirates, Philippines, India, and Oman are characterised by very high search interest in these projects (index > 50).

Within each country, a preference is observable towards either project, which can be summarised as a divergence index. For example, in the Philippines is 5 for OpenStreetMap and 58 for Wikimapia. Considering the total volume for the country (summed indices, 63 in this case), 7.9% of searches targeted OpenStreetMap (*O*), and 92.1% focussed on Wikimapia (*W*). Hence, we calculate the divergence index as *O*-*W*, so, in this case, it would be -84.2%. It is important to note that the divergence index, ranging from -100 to 100, is relative to the overall interest in the two projects, shown in Figure 5. For example, the two indices for India amount to 61 (58 for Wikimapia, 3 for OpenStreetMap), while the United Kingdom has only 7 (2 for Wikimapia, 5 for OpenStreetMap), resulting respectively in divergence indices -90 and 42. The map in Figure 6[35] shows this summary index per country, spanning from -100 (all searches for Wikimapia, represented in red) to 100 (all searches for OpenStreetMap, represented in green).

This data suggests that Wikimapia is relatively more popular in Northern Africa, Middle East, India, Southeast Asia, Latin America, Russia, and parts of Eastern Europe. By contrast, OpenStreetMap is overwhelmingly more searched for in Western Europe, North America, South Africa, Oceania, Japan, South Korea, and China. Not enough data is available for Sub-Saharan Africa and Central Asia. Interestingly, this spatial divergence retraces the digital divide between the Global North of high-income countries and the Global South, with medium- and low-income countries (Ballatore et al., 2017).

Several explanations can be advanced in this context. First, Wikimapia originated from Russia, and might have been able to produce data about Russia and neighbouring countries more effectively. OpenStreetMap started in the UK and spread primarily in Northern and Central Europe, particularly in German-speaking countries. Second, being focussed on place names and locations, Wikimapia might represent places neglected in major global products such as Google Maps, as well as projects like OpenStreetMap and GeoNames, which are deeply biased in favour of the Global North (Acheson et al., 2017): This would explain the higher volume of searches across the Global South. This data also suggests that, despite its better ethical standing and richer data model, OpenStreetMap seems to fail to capture location-related searches in medium- and low-income regions, which are instead served by Wikimapia.

5. Conclusions

In this study, we have investigated several facets of Wikimapia, providing a qualitative overview of the project's inception, ethos, and development, in a comparison with the more studied OpenStreetMap. Wikimapia can be described as a crowdsourced gazetteer, mainly derived by drawing bounding boxes on Google imagery. The data model is simpler than OpenStreetMap, but contains longer descriptive information and images, while being spatially less accurate. In

[35] Data based on query https://trends.google.com/trends/explore?date=all&q=Wikimapia,OpenStreetMap (accessed on 17 June 2017)

this sense, Wikimapia data seems more suitable for information retrieval and search tasks, rather than for the kind of spatial analysis that is often performed with OpenStreetMap vector data. This is consistent with the Google-driven business model adopted by the company, which relies on making Wikimapia places visible in Google search results to sell online advertising and to attract new users.

As expected, Wikimapia activity is characterised by contribution inequality, with few contributors producing extraordinary amounts of information. The project manages the quality of contributions with a peer-review model, heavily based on gamification and user rewards. A grid-based quality control tool obtained mixed results in promoting data completeness. According to a former developer (Milevski, 2016), a core aspect of the data production and search is the findability of small businesses. While this study does not provide a quantitative assessment of Wikimapia's data quality, we hope that this overview will be useful to ground further research.

A major challenge to Wikimapia is its uncertain standing with respect to the copyright of its content. Despite having adopted a Creative Commons license, the bulk of its vector data was derived from Google satellite imagery without explicit agreements.[36] Being for-profit and advertisement-driven, Wikimapia is also a notable exception to a commons-based ecosystem, which raises ethical questions about unpaid digital labour (Graham & Anwar, 2018) and makes the re-use of its data potentially problematic for practitioners.

Relying on behavioural data from Alexa and Google Trends, we explored the online visibility of the two projects. Wikimapia's development and usage have been deeply linked to SEO. For this reason, its popularity in Google searches spiked at the very beginning of the project in 2006. At its peak in 2009–2010, the project managed to attract up to 1M unique visitors on a daily basis (Milevski, 2016). However, searches for Wikimapia have been declining since 2012, and no major platform updates have been performed since then. Furthermore, to date, the platform's API is malfunctioning, and the project team seems inactive. By contrast, OpenStreetMap had a slow start in 2004, and overtook Wikimapia's relative search interest only in 2014. Our analysis shows that searches for these VGI projects are strongly seasonal, closely following holidays patterns in the Northern hemisphere.

Our study shows how, rather surprisingly, Wikimapia and OpenStreetMap have widely divergent and segregated geographies of searches and, very probably, use. Adopting Google search data as a proxy for popularity, we showed how Wikimapia tends to be more searched for in the Global South, particularly in Northern Africa, Middle East, India, and Russia. OpenStreetMap, on the contrary, is more visible across the Global North in most high-income countries, particularly in German-speaking parts of Europe. This might indicate that Wikimapia better represents places relatively under-represented in global, corporate web maps, as well as by projects like Wikipedia, OpenStreetMap and GeoNames. Indeed, OpenStreetMap-based humanitarian mapping represents a notable exception to this trend, as it tends to target disadvantaged, low-income areas.

[36] http://wiki.openstreetmap.org/wiki/Wikimapia#License_.28and_imagery_derived_data_issue.29 (accessed on 15 August 2017)

It is important to note that, overall, the online visibility of VGI projects, such as Wikimapia and OpenStreetMap, appears negligible compared with digital corporate players like Google and Microsoft. While Wikipedia managed to displace Britannica and Encarta, such crowdsourced mapping projects represent at most a complementary data source, and are currently very far from representing a real challenge to Google Maps. This state of affairs can be attributed to many factors, starting from limited routing capabilities, lack of highly usable and reliable mobile apps, as well as low search engine visibility for place searches.

Many questions about Wikimapia and, more broadly, about VGI are worth pursuing in future research. To understand the nature of its spatial data, quantitative studies should focus on the data, at the micro- and meso-scale, evaluating its thematic and geometric aspects in comparison with OpenStreetMap and traditional sources. More qualitative studies can target semantic, cultural, political, and sentic aspects of Wikimapia data. The truth in this critical claim about Wikimapia found on Wikipedia remains to be investigated: "map coverage is generally uneven, with some areas, usually in developing countries, being cluttered with crude outlines, private residences, subjective evaluations or advertisements".[37]

From the perspective of VGI geographies, more indicators are needed to refine our initial geography of Wikimapia and OpenStreetMap visibility, going beyond proxy measures based on search patterns. The Wikimapia community can be studied through quantitative analysis from the perspective of its demography and geography. To further enrich our VGI inventory, more sources and projects should be investigated in terms of online visibility and search patterns. Based on this study, we believe that online visibility and search engine optimisation should be of utmost concern for project owners and practitioners, impacting on the longevity and sustainability of projects. How to make VGI projects more visible for mass consumption online in absence of substantial capital is likely to remain an open question for the next ten years.

Acknowledgements

The authors gratefully acknowledge the funding from the E.U. COST Action IC1203 (European Network Exploring Research into Geospatial Information Crowdsourcing: software and methodologies for harnessing geographic information from the crowd), which provided the opportunity of identifying and discussing the research gaps on Wikimapia. They also thank Alexander Milevski, Alexa Internet, Google Trends, *naturalearthdata.com*, and Wikimapia and OpenStreetMap contributors for providing useful information, and Alexander Zipf for his insightful comments.

[37] https://en.wikipedia.org/wiki/Wikimapia#Quality_of_contents (accessed on 27 July 2017)

References

Acheson, E., De Sabbata, S., & Purves, R. S. (2017). A quantitative analysis of global gazetteers: Patterns of coverage for common feature types. Computers, Environment and Urban Systems, 64, 309–320. http://doi.org/10.1016/j.compenvurbsys.2017.03.007

Ahmouda, A., & Hochmair, H. H. (2017). Using Volunteered Geographic Information to measure name changes of artificial geographical features as a result of political changes: a Libya case study. GeoJournal, 1–19. In press.

Ballatore, A. (2014). Defacing the map: Cartographic vandalism in the digital commons. The Cartographic Journal, 51(3), 214–224. http://doi.org/10.1179/1743277414Y.0000000085

Ballatore, A., & Mooney, P. (2015). Conceptualising the geographic world: The dimensions of negotiation in crowdsourced cartography. International Journal of Geographical Information Science, 29(12), 2310–2327. http://doi.org/10.1080/13658816.2015.1076825

Ballatore, A., & Zipf, A. (2015). A Conceptual Quality Framework for Volunteered Geographic Information. In S. I. Fabrikant & M. Raubal (Eds.), Spatial Information Theory (Vol. 9368, pp. 89–107). Berlin: Springer. http://doi.org/10.1007/978-3-319-23374-1

Ballatore, A., Graham, M., & Sen, S. (2017). Digital Hegemonies: The Localness of Search Engine Results. Annals of the American Association of Geographers, 107(5), 1194–1215. http://doi.org/10.1080/24694452.2017.1308240

Bittner, C. (2016). Diversity in volunteered geographic information: comparing OpenStreetMap and Wikimapia in Jerusalem. GeoJournal. http://doi.org/10.1007/s10708-016-9721-3

Brabham, D. C. (2013). Crowdsourcing. Cambridge, MA: MIT Press.

Budhathoki, N. R., & Haythornthwaite, C. (2013). Motivation for Open Collaboration. American Behavioral Scientist, 57(5), 548–575. http://doi.org/10.1177/0002764212469364

Choi, H. and Varian, H. (2012). Predicting the Present with Google Trends. Economic Record, 88(1):2–9.

Dodge, M., & Kitchin, R. (2013). Mapping Experience: Crowdsourced Cartography. Environment and Planning A, 45(1), 19–36.

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. Annals of the Association of American Geographers, 102(3), 571–590.

Glasze, G., & Perkins, C. (2015). Social and political dimensions of the OpenStreetMap project: Towards a critical geographical research agenda. In OpenStreetMap in GIScience: Experiences, Research, and Applications (pp. 143-166). J. Jokar Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Berlin: Springer.

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

Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. Spatial Statistics, 1, 110–120. http://doi.org/10.1016/j.spasta.2012.03.002

Graham, M., & Anwar, M. A. (2018). Digital Labour. In J. Ash, R. Kitchin, & A. Leszczynski (Eds.), Digital Geographies. London, UK: Sage Publications.

Graham, M., & De Sabbata, S. (2015). Mapping information wealth and poverty: the geography of gazetteers. Environment and Planning A, 47(6), 1254–1264. http://doi.org/10.1177/0308518X15594899

Graham, M., De Sabbata, S., & Zook, M. A. (2015b). Towards a study of information geographies: (im)mutable augmentations and a mapping of the geographies of information. Geo: Geography and Environment, 2(1), 88–105. http://doi.org/10.1002/geo2.8

Graham, M., Straumann, R. K., & Hogan, B. (2015a). Digital divisions of labor and informational magnetism: Mapping participation in Wikipedia. Annals of the Association of American Geographers, 105(6), 1158–1178. http://doi.org/10.1080/00045608.2015.1072791

Janc, K. (2016). A Global Approach to the Spatial Diversity and Dynamics of Internet Domains. Geographical Review, 106(4), 567-587.

Jokar Arsanjani, J., Mooney, P., Helbich, M., Zipf, A. (2015): An exploration of future patterns of the contributions to OpenStreetMap and development of a Contribution Index, Transactions in GIS, 19(6): 896–914. DOI: 10.1111/tgis.12139.

Jokar Arsanjani, J., Mooney, P., Zipf, A., Schauss, A., (2015): Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets. In: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., (eds) OpenStreetMap in GIScience: experiences, research, applications. ISBN:978-3-319-14279-1, pp. 37-51, Berlin: Springer.

Kalantari, M., & La, V. (2015). Assessing OpenStreetMap as an open property map. In: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., (eds) OpenStreetMap in GIScience: experiences, research, applications (pp. 255-272). Berlin: Springer.

Kuhn, W. (2012). Core concepts of spatial information for transdisciplinary research. International Journal of Geographical Information Science, 26(12), 2267–2276.

Milevski, A. (2016, 14 June) Skype interview.

Mosco, V. (2004). The Digital Sublime: Myth, Power, and Cyberspace. Cambridge, Massachusetts: MIT Press.

Neis, P., Zielstra, D. & Zipf, A. (2013). Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. Future Internet. 5(2). pp. 282-300. doi:10.3390/fi5020282

Palen, L., Soden, R., Anderson, T. J., & Barrenechea, M. (2015). Success & scale in a data-producing organization: The socio-technical evolution of OpenStreetMap in response to humanitarian events. In Proceedings of the 33rd annual ACM conference on human factors in computing systems (pp. 4113-4122). New York: ACM.

Preis, T., Moat, H. S., Stanley, H. E., & Bishop, S. R. (2012). Quantifying the advantage of looking forward. Scientific Reports, 2, 350. http://doi.org/10.1038/srep00350

See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., … Rutzinger, M. (2016). Crowdsourcing, Citizen Science or Volunteered Geographic Information? The Current State of Crowdsourced Geographic Information. ISPRS International Journal of Geo-Information, 5(55), 1–23. http://doi.org/10.3390/ijgi5050055

Senaratne, H., Mobasheri, A., Ali, A. L., Capineri, C., & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. International Journal of Geographical Information Science, 31(1), 139–167. http://doi.org/10.1080/13658816.2016.1189556

Stein, K., D. Kremer, and C. Schlieder (2014). Spatial Collaboration Networks of OpenStreetMap. In: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., (eds) OpenStreetMap in GIScience: experiences, research, applications (pp. 167–186). Berlin: Springer.

Stevens, M., Vitos, M., Altenbuchner, J., Conquest, G., Lewis, J., & Haklay, M. (2014). Taking participatory citizen science to extremes. IEEE Pervasive Computing, 13(2), 20–29. http://doi.org/10.1109/MPRV.2014.37

Stephens-Davidowitz, S. I. (2013). Essays using Google Data. PhD Dissertation. Cambridge,MA: Harvard University.