REVISTA DO DEPARTAMENTO DE GEOGRAFIA
Universidade de São Paulo
www.revistas.usp.br/rdg - ISSN 2236-2878
Volume 43 (2023), e181687
DOI: 10.11606/eISSN.2236-2878.rdg.2023.181687

Volunteered Geographic Information: a 10-year bibliometric investigation

Informação Geográfica Voluntária: uma investigação bibliométrica de 10 anos

Tatiane Ferreira Olivatto

Departamento de Engenharia Civil, Universidade Federal de São Carlos, São Carlos, SP, Brasil.
E-mail para correspondência: tatianeolivatto@yahoo.com.br

Recebido (*Received*): 08/02/2021
Aceito (*Accepted*): 20/03/2023

Resumo: A Informação Geográfica Voluntária (VGI) tornou-se mais evidente ao mesmo tempo em que as plataformas de código aberto se tornaram populares em todo o mundo, ambas resultantes do fácil acesso das pessoas às informações geográficas em seus smartphones. Com o objetivo de investigar os principais aspectos deste campo de pesquisa, foi desenvolvida uma investigação bibliométrica com foco num período de 10 anos (2011-2020). A análise foi realizada com base no banco de dados Scopus e nos *softwares* VOS Viewer e Bibliometrix, abordando: publicações ao longo dos anos, tipos de documentos, campos de estudo, principais periódicos, principais artigos, países, autores e palavras-chave mais recorrentes. Os resultados iniciais indicaram que: as publicações aumentaram a uma taxa anual de 21.69%, o tipo de documento mais publicado foi artigo e apenas 16 periódicos foram responsáveis por 33.33% dos 1200 artigos publicados. EUA, Alemanha e Reino Unido são os principais países que pesquisam VGI e os dois últimos também são países-sede dos principais autores. Apesar do termo VGI ter sido definido em meio a Ciência Cidadã, a rede de ocorrência de palavras-chave mostrou que SIG (Sistema de Informação Geográfica) é um campo de estudo de destaque. Contudo, a rede de visualização com base em média de publicações por ano revelou a Ciência Cidadã como um campo de pesquisa ainda em avanço. Palavras-chave como OpenStreetMap, qualidade dos dados, avaliação da precisão, mídias sociais e coletividade mostraram-se mais difundidas no campo, o oposto ocorre com aplicações em áreas urbanas, uso do solo e serviços ecossistêmicos. No geral, os indicadores bibliométricos revelaram-se eficazes para acessar a VGI como tópico de pesquisa e indicaram uma tendência promissora em temas envolvendo redes sociais, sensoriamento remoto, área urbana, colaboração coletiva e PPGIS.

Palavras-chave: Informação Geográfica Coletiva; Mapeamento Colaborativo; Mapeamento Participatório; Bibliometria.

Abstract: *Volunteered Geographic Information (VGI) has become more evident at the same time as open-source platforms become worldwide popular, both resulting from people easily accessing geographic information on their smartphones. Aiming to investigate the main aspects of this research field, a bibliometric investigation was developed focusing on 10-year period (2011-2020). The analyses were performed based on Scopus database, VOS Viewer and Bibliometrix softwares, approaching: publications over years, document types, subject areas, core sources, main papers, countries, authors and most recurrent keywords. The initial results indicated that: publications have increased at an annual rate of 21.69%, the most published document type was article and only 16 journals were responsible for 33.33% of those 1200 articles published. USA, Germany and UK are major countries researching VGI and the last two are also host countries of the main authors. Although the term VGI has been defined among Citizen Science, the network of keywords occurrence showed that GIS (Geographic Information Systems) is an outstanding study field. However, the network visualization based on average publication per year revealed Citizen Science as a research field still moving forward. Keywords such as OpenStreetMap, data quality, accuracy assessment, social media and crowdsourcing showed to be more widespread among the field, the opposite occurs with applications in urban areas, land use and ecosystem services. Overall, the bibliometric indicators have revealed to be effective in order to access VGI as a research topic and indicated a promising trend in themes involving social media, remote sensing, urban area, crowdsourcing and PPGIS.*

Keywords: *Crowdsourced Geographic Information; Collaborative Mapping; Participatory Mapping; Bibliometrics.*

1. Introduction

Volunteered Geographic Information (VGI) has emerged mainly in the last 15 years, since a great amount of people has access to devices supplied with GPS sensors, like smartphones and personal computers, being able to consume and produce geographic information (CAPINERI *et al*., 2016).

The consumption relationship can be easily identified when people search for an address in Google Maps or plan a route in Waze. The production relationship can be harder to diagnose because, on some occasions, occurs indirectly, for instance, via Facebook, Twitter, Flickr and Instagram through geotagging. On the other hand, the production relationship occurs directly when users approach a specific platform to contribute, for example, developing a Wikimapia webpage or reporting a car accident on Waze. As a result of these characteristics, VGI has been mostly studied in Citizen Science (HAKLAY, 2013).

In the context of Citizen Science, Michael Goodchild coined the term Volunteered Geographic Information in 2007 as "volunteers who collect geographic data in a collaborative way, allocating citizens as sensors" (GOODCHILD, 2007).

Despite the terminology was framed in recent years, publications about VGI – even adopting a different nomenclature or no term at all - can be founded since ancient times. Some authors consider maritime navigation maps, from 1500, an evidence of old collaborative mapping because they were developed with contributions from sailors returning from their voyages. Goodchild (2007) described this kind of volunteers: "they are largely untrained and their actions are almost always voluntary, and the results may or may not be accurate. But collectively, they represent a dramatic innovation that certainly had profound impacts on Geographic Information Systems (GIS) at the time".

Even mentioning VGI during ancient times, the same author refers specifically to VGI as a web phenomenon, therefore, a recent phenomenon. Goodchild (2007) also specifies that VGI is user-driven, meaning users who are predisposed to map a certain region. Haklay (2013) goes further by trying to understand the role of the volunteer more deeply. He approaches this aspect from the point of view of Geographical Citizen Science. According to the author, volunteers fit as active or passive.

An active volunteer is aware of your collaboration, for instance, when adding a street or a building at an open mapping platform or sending a geotagged photo to a project's hub. On the other hand, a passive volunteer collaboration occurs without active participation (HAKLAY, 2013), being one of the most common situations when mobile users activate the GPS function while using a routing app, allowing real-time traffic data collection.

Currently, the most widespread active VGI project is OpenStreetMap (OSM), connecting more than 6.4 million registered users who contribute to the project at some level, including about 7.8 billion uploaded GPS points, 6.0 billion nodes, 600 million ways and 7.0 million tags and relations (OSM, 2020).

In the literature, studies related to VGI - also referred as crowdsourcing geographic information, collaborative mapping or even participatory mapping - are frequently mentioned along with OSM researches, focusing on data quality, accuracy, credibility, completeness and applications (MONDZECH and SESTER, 2011; CASTRO *et al*., 2019). Researches approaching data completeness and quality are considered essential to describe data usability (HAKLAY, 2013). In this context, positional accuracy, geometric feature, absolute measures and tags existence are largely investigated (MONDZECH and SESTER, 2011). There are also some quality open-source tools for data quality measurement such as OSM Quality Assurance Editor (ZIPF *et al*., 2016).

Furthermore, there are several studies related to VGI applications, most based on OSM. Examples are tools for disaster risk management, as used in 2010 after Haiti earthquake (DE ALBUQUERQUE *et al*., 2016); for urban resilience, as flooding situation awareness, geophysical hazards and global environmental change monitoring (PAUL *et al*., 2019). Researches dealing with routing services are frequently mentioned closely to VGI studies, for example, Bexhill OSM for pedestrians routing; FacilMap for vehicles, cyclists and pedestrians custom routing; OpenRouteService, Routino and OpenTripPlanner for accessible routing (SCHMITZ *et al*., 2008; NEIS and ZIELSTRA, 2014; ZIPF *et al*., 2016).

There are also some innovative efforts to investigate the VGI relationship with other areas, for example, tourism (CHEN *et al*., 2018; DEVKOTA *et al*., 2019; FERREIRA, 2019), quality of life for citizens during urban transit (KELER and MAZIMPAKA, 2016; NOVACK *et al*., 2018), remote sensing, land use planning and smart cities (CHEN *et al*., 2019; DORNER *et al*., 2019; MIAO *et al*., 2019; SENTURK and KEBE, 2019). Also, there are state of art methods in big data and artificial intelligence being developed concomitantly to VGI applications (ZIPF *et al*., 2016; FERREIRA, 2019).

Overall, VGI has shown potential to be a significant source of geographic information, supporting a range of fields among society. Another advantage of VGI consists in reducing prices in data collection by motivating individuals to act voluntarily. On the other hand, by acting voluntarily, quality of information can be questionable (GOODCHILD, 2007), consequently, the difficulty in measuring the quality of the information is often pointed out as one of the weaknesses of VGI.

All these acknowledged, one of the most important outputs from VGI may be a change in geographic information characteristics, mainly in its creation, sharing, dissemination and application (ELWOOD *et al*., 2012), thus, drawing attention to the need for understanding VGI as a research field. According to Capineri *et al*. (2016), articles matching "volunteered geographic information" totalized 83 on Google Scholar in 2007, the same year that the terminology emerged. From 2009 to 2014, the same searching increased from less than 200 to almost 1000 articles.

In line with the growing volume of scientific production, the topics related to VGI are several, including theoretical conception, types of collaboration, aspects of quality, accuracy, precision, mapping projects, completeness and reliability of information. In addition, there is a segment of publications focusing on applications, such as disaster management, land use, tourism, environmental monitoring, accessibility, routing, etc (CAPINERI *et al*., 2016; PAUL *et al*., 2019).

All facts described bring about a strong segregation of academic publications and make the evaluation of the study field progress harder. In this context, the development of bibliometric indicators can be a useful tool in order to quantify academic performance, identify relevant topics of a given science and analyse its evolution over time (SANTIN *et al*., 2019).

Thus, this work presents a bibliometric investigation regarding VGI research topic focusing on the 10-year period from 2011 to 2020. The aspects analysed were: number of publications over the years, most predominant document type, subject areas, core sources, most cited publications, main countries committed with the topic, most influent authors and frequent employed keywords.

2. Methodology

The methodological procedures were developed based on Scopus, an online version scientific citation index database (ELSEVIER, 2021). This database provides a comprehensive citation search, including others databases, allowing multidisciplinary research with international coverage (ROMANELLI *et al*., 2018). Each step followed during methodology development is summarized in **Figure 1**.

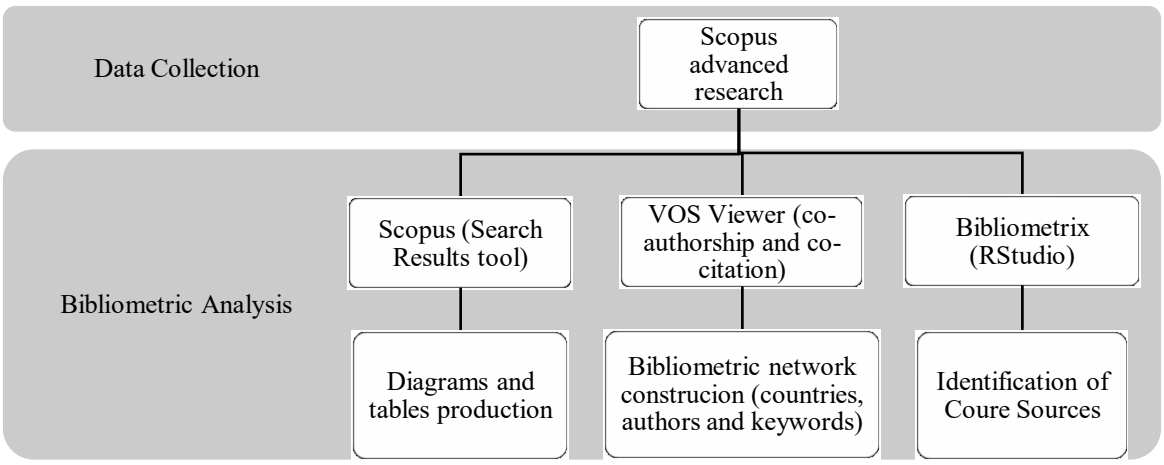


Figure 1: Methodological flowchart.

2.1. Data collection

Although the central topic investigated is VGI, in order to cover the entire field of study, common synonyms were also submitted to Scopus search database - being titles, abstracts and keywords investigated. Search date was February 05 2021. The searching criterion and specific terms inserted, following by settings and subsequent searching returns are showed in **Table 1**.

Table 1: Searching criteria.

Keywords parameters as inserted in Scopus search engine: TS (topic) = ("volunteer* geographic information" OR "crowdsourc* geographic information" OR "collaborative mapping" OR "participatory mapping")		
1995-2021	2011-2020	2011-2020
2208 results	1996 results	1200 results
	Refined by: period = from 2011 to 2020	Refined by: type of document = article

As can be seen in **Table 1**, keywords parameters inserted in Scopus search engine have some specificities. Firstly, asterisks (*) are used to include words with different endings for a same stem, as in plurals, for example. Second, quotation marks (" ") are useful to find compound terms. Finally, Boolean operators connect the search words together to either narrow or broaden a set of results. Taking advantage of these search operators, besides "volunteer[ed/ing] geographic information", the synonyms identified in introduction section – and its derivations - were included: "crowdsource[ing/ed] geographic information", "collaborative mapping" and "participatory mapping".

From 1995, the oldest record, to 2021, search date, there are 2208 records available at Scopus. The same search was performed at Web of Science, another well consolidated indexing base (WANG *et al.*, 2016), returning 1748 records. Due to the greater number of documents covered, opting for Scopus proved to be more appropriate for this research topic.

Thereafter, considering the 10-year period focus of this study, from 2011 to 2020, there are 1996 records, representing 90.40% of all publications. Thus, by investigating last 10 years, most publications were covered. In addition, aiming to ensure that the investigated information is scientifically accurate, further bibliometric analyses were restricted only to article document type, considering that they are peer reviewed.

2.2. Bibliometric analysis

The results from bibliometric search were analysed firstly throw "Analyse Search Results" tool available at Scopus website and then exported to bibliometric softwares such VOS Viewer and Bibliometrix.

The Scopus Analyse Search Results tool allows users to extract data values from a specific field of interest - as publication years, document types, authors, Scopus subject areas, countries, language, including others - and then visualize the results in ranked order (ELSEVIER, 2021). The resulting data can be visualized at the online tool itself, throw diagrams and tables, or can be downloaded as ".csv" format for later handling.

The VOS Viewer is a free software towards bibliometric networks building and visualizing. The resulting network can be based in several parameters such as co-authorship (of authors, organizations or countries) and co-occurrence (of authors keywords) (VAN ECK and WALTMAN, 2014).

The bibliometric analysis using VOS Viewer software provides visualization of bibliometric networks, also known as 'science mapping'. In order to interpret bibliometric networks, some aspects of them need to be understand: (i) node, (ii) edges/link and (iii) cluster.

According to Van Eck and Waltman (2014), a bibliometric network consists of nodes and edges. The nodes represent the investigated items, for instance authors, journals, organizations, countries or keywords. The relationship between nodes is represented by edges. The size of nodes refers to how often the term appears and edge width indicates the intensity of correlation. The VOS Viewer classifies the nodes of a network into clusters, being one node belonging to only one cluster. Each cluster contains nodes closely related and they are represented in the visualization of a bibliometric network by different colors (VAN ECK and WALTMAN, 2014).

Overall, the main characteristic of Scopus tools is the quantitative nature of analysis. While VOS Viewer allows also qualitative analysis, enabling further investigations. For this paper development, the survey fields at Scopus Analyse Search Results tool were publications years, document type, subject areas, source titles and main papers. At VOS Viewer software, the topics investigated were countries, authors and keywords.

Core sources analysis was performed at Biblimetrix in RStudio. The methodology to identify core sources is based on Bradford's Law of diminishing returns and scattering: *"there are a few very productive periodicals, a larger number of more moderate producers, and a still larger number of constantly diminishing productivity"* (BRADFORD *et al*., 1953). Bradford *et al*. (1953) claimed that a few number of journals can cover the most important and fundamental literature of a field. By using this law, the scholarly journals were separated into two zones in terms of number of publications, wherein the journals comprised in the first zone refers to core sources in the field.

3. Results and discussion

3.1. Bibliometric analysis at Scopus and Bibliometrix

The survey on volunteered geographic information - and its synonyms - from 2011 to 2020 returned a total of 1996 documents published as articles, proceedings papers, book chapters, reviews, conference reviews and other types of documents. The overall distribution of document types can be seen in **Figure 2**.

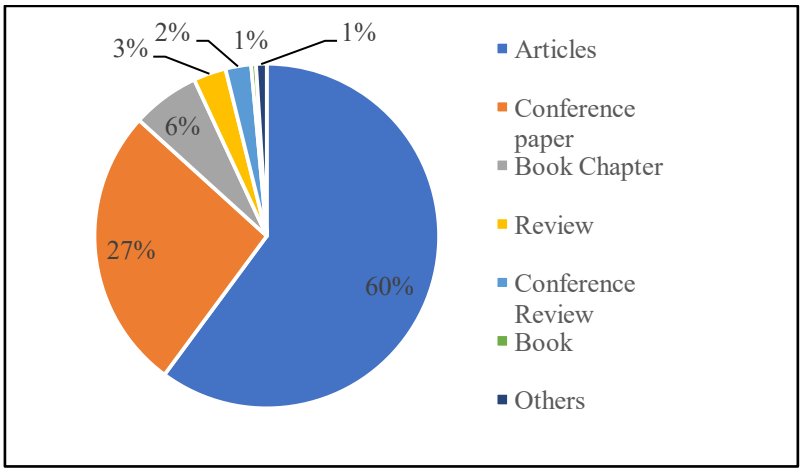
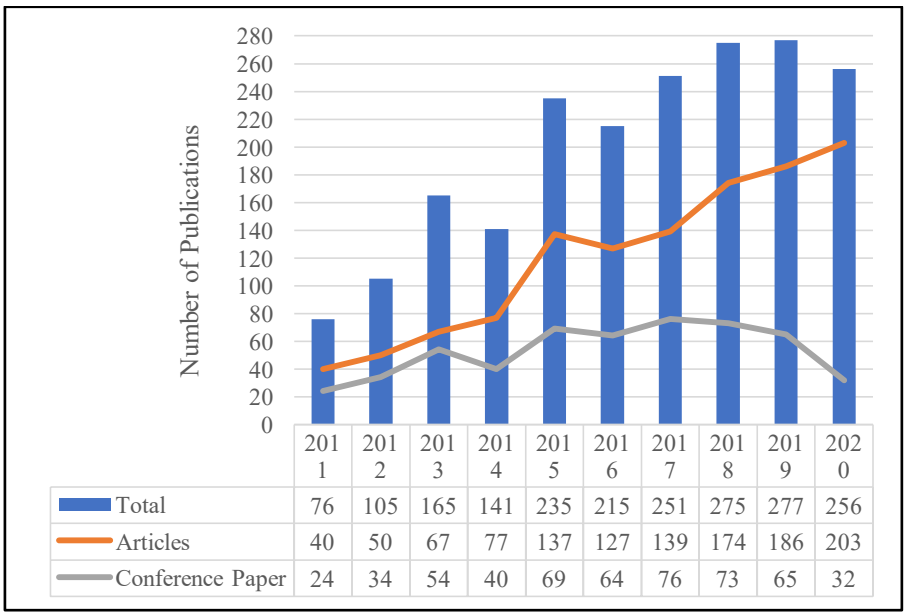


Figure 2: Document Types. Note: "Others" category includes editorial, data paper, erratum, note, letter, short survey and undefined.

Among document types returned during the bibliometric search, articles are majority, 60% of publications for the analysed period, followed by conference papers, with 27%. The predominance of articles repeats over the years, as showed in **Figure 3** (red line).



	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Total	76	105	165	141	235	215	251	275	277	256
Articles	40	50	67	77	137	127	139	174	186	203
Conference Paper	24	34	54	40	69	64	76	73	65	32

Figure 3: Total publications from 2011 to 2020 (blue bars), emphasis for articles (red line) and conference papers (green line).

From **Figure 3** is also possible to realize that the volume of scientific production increased over the years reaching a peak in 2019 (see blue bars). Articles showed a more ascendant curve, at an annual average growth of 21.69% in the 10-year period. This reflects an overall ascent trend of this research field, however, conference papers (green line) showed a slightly decrease from 2017 onwards, decreasing drastically in 2020 - that is likely to be related to the slowdown in academic events and conferences due to COVID-19 pandemic. On the other hand, articles remained on the rise throughout the period.

Among subject areas to which the article belongs, six are predominant: Social Sciences, Earth Planetary Sciences, Environmental Science, Agricultural and Biological Sciences, Computer Science and Engineering. Some of them appears to be more related to theoretical and empirical research and other to application fields. The variety of research fields indicate multidisciplinarity, being Social Sciences and Earth Planetary Sciences outstanding.

Considering that most publications are articles, some embracing studies were done focusing on this kind of publication. Taking into account that articles are peer reviewed, this restriction also means a study of quality checked documents. Altogether, the 1200 papers (between 2011 and 2020) were published in 472 journals.

Using Bibliometrix, 16 core sources were identified (see **Figure 4**), where 400 of 1200 papers were published in these core sources (33.33% of them). Based on Bradford's Law, Garfield (1972) estimates "*that between 500 and 1000 journals were needed to cover 95 per cent of the significant literature published in a given field*". In this work, 113 journals were responsible for the publication of 95% of the investigated articles.

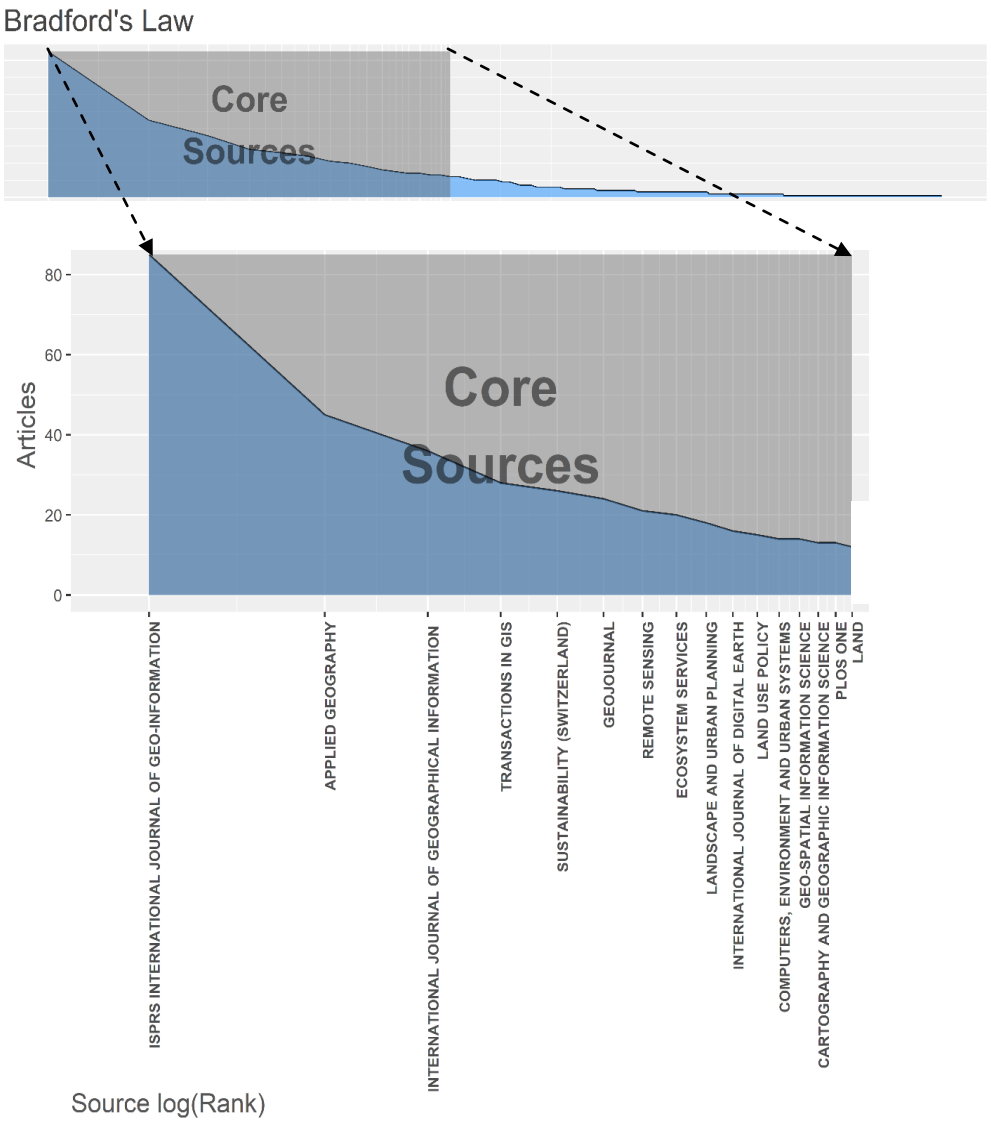


Figure 4. Core Sources identified by Bradford's Law. Built in Bibliometrix.

Another relevant aspect related to sources is the quality of them. One widespread quality of journals measurement is Citescore, allowing ranking, evaluating, categorizing and comparing journals. In simple words, Citescore is a Scopus bibliometric method to evaluate the importance of scientific journals in their respective areas (MENDES and MARZIALE, 2002).

In mathematical terms, "*Citescore is based on the number of citations to* [peer-reviewed] *documents (articles, reviews, conference papers, book chapters, and data papers) by a journal over four years, divided by the number of the same document types indexed in Scopus and published in those same four years*" (ELSEVIER, 2021). **Table 2** is ordered by Citescore (from highest to lowest) and shows, respectively, source titles (Journals), Citescore, total citation and ranking number related to the core sources identified in **Figure 4**.

The Citescore average of these journals is 6.7, being 11.6 the highest and 3.0 the lowest. Considering only core sources papers, 61% of them were published in journal with Citescore above the average 6.7, indicating a high level of research quality. ISPRS International Journal of Geo-Information was the journal with the highest number of publications (85), however with the fifth lower Citescore between core sources. The following journals with highest number of publications are Applied Geography (45) and International Journal of Geographical Information Science (36), both with Citescore over the average, 7.2 and 8.2, respectively. Regarding to the journal with the highest Citescore (11.6), 18 papers were published in it.

As already mentioned, Citescore is a measure of the frequency with which the average document has been cited in a particular period. Based on this premise and aiming to verify the importance of this indicator, bibliometric information of the 15 most cited papers were collected and are summarized in **Table 3**.

A comparison between **Table 3** and **Figure 4** discloses that 8 of the 15 most cited papers were published in core sources identified by Bradford's Law. As showed in the last column of Table 3, the main core source (R 1) published the tenth (ID X) and twelfth (ID XII) most cited papers. In a similar analysis, the most cited paper (ID I) was published in the eleventh core source (R 11).

According to Okubo (1997), there are two main reasons that increase the probability of papers to be cited: the longest time indexed and a high level of innovation or new approaches in the area. In this sense, some analysis based on the information available in **Table 3**, specifically in column "year", could help to understand these reasons.

Table 2: Core sources ranked by Citescore

Source Title	CS 2020	TP	R
LANDSCAPE AND URBAN PLANNING	11.6	18	9
ECOSYSTEM SERVICES	11.2	20	8
COMPUTERS, ENVIRONMENT AND URBAN SYSTEMS	10.1	14	12
INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	8.2	36	3
LAND USE POLICY	7.5	15	11
GEO-SPATIAL INFORMATION SCIENCE	7.4	14	13
APPLIED GEOGRAPHY	7.2	45	2
REMOTE SENSING	6.6	21	7
INTERNATIONAL JOURNAL OF DIGITAL EARTH	6.6	16	10
CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE	5.7	13	14
PLOS ONE	5.3	13	15
ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION	4.6	85	1
TRANSACTIONS IN GIS	4.6	28	4
SUSTAINABILITY (SWITZERLAND)	3.9	26	5
GEOJOURNAL	3.0	24	6
LAND	3.0	12	16

CS 2020 = Citescore 2020 (from Scopus Sources Search); TP = Total Publication, R = Ranking number related to core sources identified in **Figure 4** (the first positions in the ranking refers to journals with the largest number of publications).

Table 3: Most cited papers.

ID	Authors	Title	Year	Source Title	NC	R
I	Plieninger T., Dijks S., Oteros-Rozas E., Bieling C.	Assessing, mapping, and quantifying cultural ecosystem services at community level	2013	LAND USE POLICY	529	11
II	Goodchild M.F., Li L.	Assuring the quality of volunteered geographic information	2012	SPATIAL STATISTICS	427	Ø
III	Elwood S., Goodchild M.F., Sui D.Z.	Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice	2012	ANNALS OF THE ASSOCIATION OF AMERICAN GEOGRAPHERS	364	Ø
IV	Fagerholm N., Käyhkö N., Ndumbaro F., Khamis M.	Community stakeholders' knowledge in landscape assessments - Mapping indicators for landscape services	2012	ECOLOGICAL INDICATORS	246	Ø
V	Palomo I., Martín-López B., Potschin M., Haines-Young R., Montes C.	National Parks, buffer zones and surrounding lands: Mapping ecosystem service flows	2013	ECOSYSTEM SERVICES	232	8
VI	Michael N., Shen S., Mohta K., Mulgaonkar Y., Kumar V., Nagatani K., Okada Y., Kiribayashi S., Otake K., Yoshida K., Ohno K., Takeuchi E., Tadokoro S.	Collaborative mapping of an earthquake-damaged building via ground and aerial robots	2012	JOURNAL OF FIELD ROBOTICS	230	Ø
VII	Middleton S.E., Middleton L., Modafferi S.	Real-time crisis mapping of natural disasters using social media	2014	IEEE INTELLIGENT SYSTEMS	209	Ø
VIII	Van Berkel D.B., Verburg P.H.	Spatial quantification and valuation of cultural ecosystem services in an agricultural landscape	2014	ECOLOGICAL INDICATORS	208	Ø
IX	Stefanidis A., Crooks A., Radzikowski J.	Harvesting ambient geospatial information from social media feeds	2013	GEOJOURNAL	198	6
X	See L., Mooney P., Foody G., Bastin L., Comber A., Estima J., Fritz S., Kerle N., Jiang B., Laakso M., Liu H.-Y., Milèinski G., Nikšieč M., Painho M., Podör A., Olteanu-Raimond A.-M.-R., Rutzinger M.	Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information	2016	ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION	192	1
XI	de Albuquerque J.P., Herfort B., Brenning A., Zipf A.	A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management	2015	INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE	183	3
XII	Neis P., Zipf A.	Analyzing the contributor activity of a volunteered geographic information project - The case of OpenStreetMap	2012	ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION	180	1
XIII	Klain S.C., Chan K.M.A.	Navigating coastal values: Participatory mapping of ecosystem services for spatial planning	2012	ECOLOGICAL ECONOMICS	179	Ø

XIV	Barron C., Neis P., Zipf .	A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis	2014	TRANSACTIONS IN GIS	173	4
XV	Brown G., Weber D.	Public Participation GIS: A new method for national park planning	2011	LANDSCAPE AND URBAN PLANNING	154	9

R=Ranking; NC=Number of citation; R= Ranking number related to core sources identified in **Figure 4** (the first positions in the ranking refers to journals with the largest number of publications).

The six most cited papers were published in 2012 or 2013. Previous to these years, in 2010, Bing allowed the use of satellite images in OSM platform in order to support Humanitarian OSM Team (DE ALBUQUERQUE *et al*., 2016). After that, in 2012, Google Maps started to charge for some services and, consequently, OSM became more requested (PANDIAN, 2012; BUCZKOWSKI, 2012; FOSSUM, 2012). Then, the longest time indexed and the described historical facts can clarify these high numbers of citations – in addition to the evidence already stablished in introduction section that OSM is the most widespread VGI project.

The most recent paper was published in 2016, corresponding to tenth position in the ranking. Against Okubo (1997) affirmation, the reason for this paper to be cited a lot may not be time or innovation, but nature of publication: a state of art review article.

Considering that data managed directly at Scopus online tools does not allow deeply analysis, for example, regarding to other possible main topics and trends, at this point, the transition to VOS Viewer methodology seems to be justified.

3.2. Bibliometric analysis at VOS Viewer

After importing data from the Scopus database into VOS Viewer software, analysis based on the basic survey (2011-2020, 1200 articles) were performed. In order to understand international collaboration and productivity, the initial explorations at VOS Viewer focused on publication and citation amount between countries. **Figure 5** shows the bibliometric network for the most representative countries among the study field of VGI.

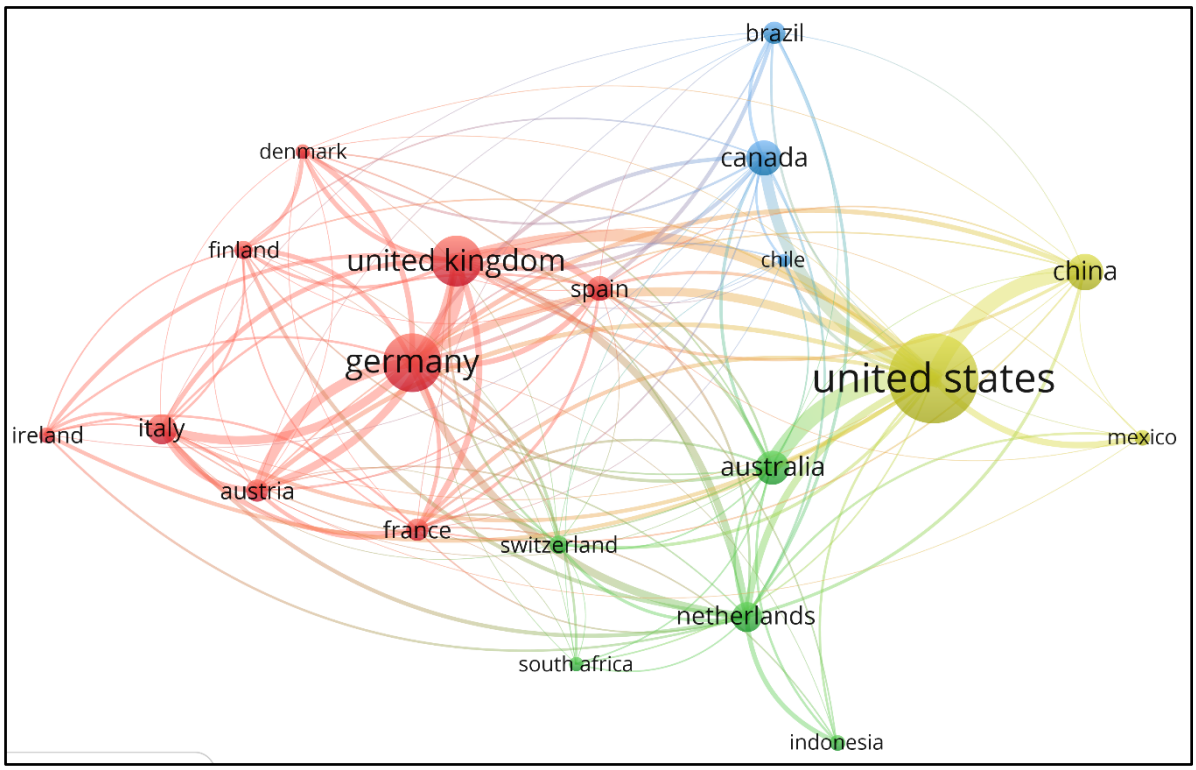


Figure 5: Co-authorship bibliometric map of countries. Built in VOS Viewer.

In **Figure 5** there are 4 clusters - see different colors - used to classify cooperation patterns between individuals from these countries. Also, the shorter line between circles the closer the academic relationship. These clusters could indicate four different scientific camps or communities on VGI research.

This network configures an arrangement where the 3 countries with greater number of publications and citations (see node size) are disposed only in 2 clusters, showing certain dominance departing from United States of America (USA), Germany and United Kingdom (UK) - from the first to third position. Despite that, there is an intense exchange of collaboration between countries (see edges width between nodes).

The biggest cluster is the red one, with 9 countries; followed by the green with 5 countries and yellow and blue with 3 countries each. South American countries, including Brazil, are all contained in blue cluster, cooperating strongly with Canada. Developing countries, such as South Africa and Indonesia, has been closely to developed countries such as Netherlands, Australia and Switzerland, as illustrated in green cluster. For developing countries, where overall currencies are less valued or even where there are limited financial resources for research, cooperation is essential in order to publish in core sources – which are mostly hosted in developed countries and, therefore, have publication fees in currencies of those countries.

Although United States is the country that most publishes in the field, cooperation between Germany and UK is larger than between them and USA (see edges width in **Figure 5**). This indicates that USA research topics/line differs in a certain way from Germany and UK. There is also a significant connection between USA and China, which also stands out on the international stage. General international cooperation is considerably miscellaneous; however, interaction inside red cluster is more intense. One aspect that could explain this is that all countries in red cluster are Europeans.

VOS Viewer was also applied to analyse cooperation between authors. **Figure 6** shows the bibliometric network for the most representative authors among the study field of VGI.

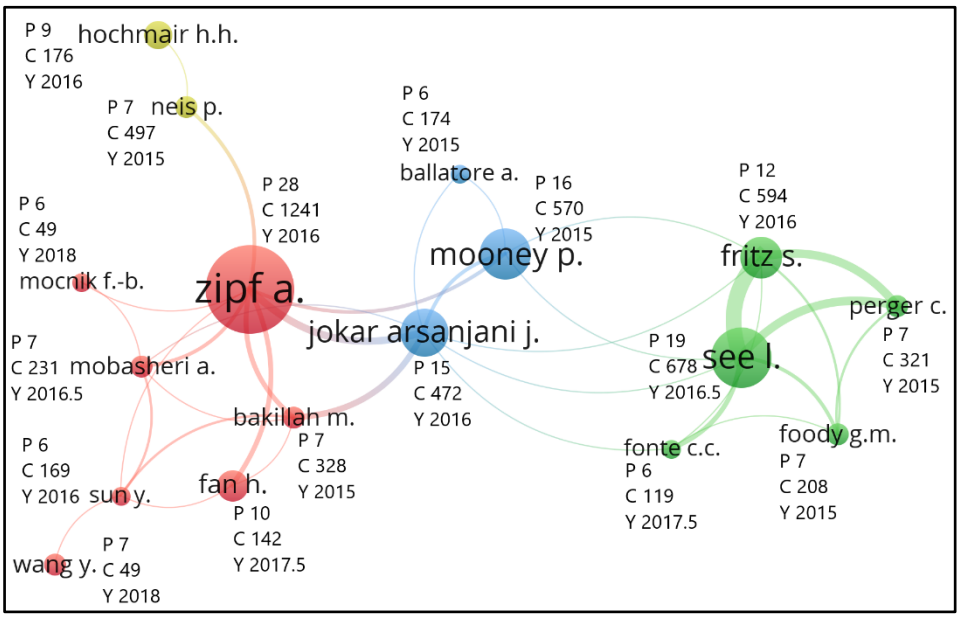


Figure 6: Co-authorship bibliometric map of authors. Built in VOS Viewer. Note: P = total papers; C = total citation; Y = average publication year.

In **Figure 6**, the size of circles represents the number of papers (the larger the circle, the more the documents) and the colors differs clusters, representing relations of collaboration - the resulting network has 4 clusters. The red cluster contains Zipf, author with the largest number of publications and citations (see circle size). The following 4 authors who have the greatest numbers of citations are See, Fritz, Mooney and Arsanjani (the first two in the green cluster and the last two in the blue cluster).

The author with the highest average citation (citation÷papers) is Neis (71.00), followed by Fritz (49.50), Zipf (44.32), See (35.68), Mooney (35.62) and Arsanjani (31.47) – except for Neis, the same as the top five in terms of total citation. The average publication year of these six authors range from 2015 to 2016.5. On the other hand, the authors with the lowest average citation are Mocnik (8.17) and Wang (7.00), both with the most recent average publication year (2018).

Cooperation between Fritz and See is the strongest, followed by the cooperation between both and Perger, all belonging to the green cluster. The link between this cluster and the blue one is provided by Mooney and Arsanjani. Also, link cooperation between Zipf and Neis are the central connection between red and yellow clusters. Cooperation inside red cluster is almost entirely directed to one author, Zipf. At the same time, there is no direct interaction between Zipf, most cited, and See, second most cited.

Figure 5 and **Figure 6** further analysis reveals a relation between countries and authors, highlighting the networks consistence. Zipf and See, at the top of authors ranking, are German and British, as well identified as main countries. Other relations were diagnosed comparing authors and countries network and **Table 3** (most cited papers). Authors contained in **Figure 6** appears 8 times in **Table 3.** In terms of countries, frequency is higher, 25 times.

In addition to these author analysis, co-citation was studied regarding to the relationship between two authors based on the number of publications in which they appear concomitantly cited (Van Eck and Waltman, 2014). **Figure 7** shows the co-citation bibliometric network of authors.

In **Figure 7**, Zipf is the second most representative author - in Figure 6 he is the first – symbolizing a high degree of scientific influence among other authors. This time, Goodchild, who coined the term VGI, is part of the network. He is also author of the second and third most cited papers (contained in **Table 3**). Other authors that differ from **Figure 6** are Brown, Haklay, Elwood, Zielstra, Graham, Corcoran, Antoniou and Martin-Lopez.

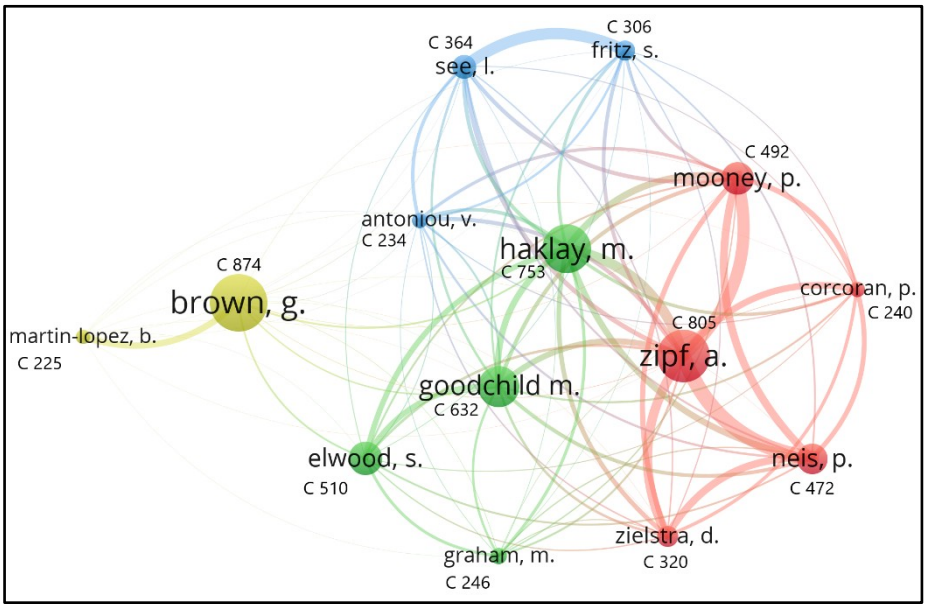


Figure 7: Co-citation bibliometric map of authors. Built in VOS Viewer. Note: C = citation.

According to Zhao and Strotmann (2008), co-citation of authors can also support investigations related to intellectual structure of influences in a scientific field as they are perceived by active authors. In this sense, by looking into Figures 6 and 7 similarities, it is possible to identify authors that are both frequently cited and infer a strong influence among VGI research topic. They are Zipf, Mooney, Neis, See and Fritz.

The next step of the proposal at VOS Viewer targeted keywords. In a first moment, the methodology demonstrated some limitations with respect to abbreviations, plurals and different ways to write the same term. For example, the terms “volunteered geographic information” and “vgi”; “geographic information system” and “gis” were counted separately. The same happened to the terms “openstreetmap”, “osm”, “open street map” and “(osm)”.

The main issue is that as they were recorded singly, they seem to appear less than other terms, when actually they are dominant keywords. To fix this limitation, the exported file in “.txt” extension was edited – using “Find and Replace” command - and the mentioned cases were rewritten and grouped together to “volunteered geographic information”, “gis” and “openstreetmap”.

Figures 8 and **9** show the resulting bibliometric network based on co-occurrence of keywords. These networks could give an indication of most frequent research themes and also identify areas or sub-areas for further development (LULEWICZ-SAS, 2017). This step of VOS Viewer methodology intended to equalize the limitation founded at the end of Scopus Analyse Search Results tool.

The network showed in **Figure 8** revealed 4 clusters (items closely related), being the red the one with the most keywords. This cluster starts with "GIS" and "mapping" as mandatory themes, evidencing publications in which VGI is approached from GIS perspective. In the same cluster, the introduction of the term "PPGIS", acronym for Public Participation Geographic Information System, add a new trend to this approach. According to Garcia *et al*. (2020) PPGIS is a "collection of methods and technologies whose main objective is the use of GIS to facilitate citizens participation in decision-making process". Brown (2017) also describe PPGIS as participatory mapping. In this sense, the inclusion of terms such as "participatory mapping", "local participation" and "decision making" in this cluster seems to be justified. "Land use" and "ecosystem service" may designate application keywords.

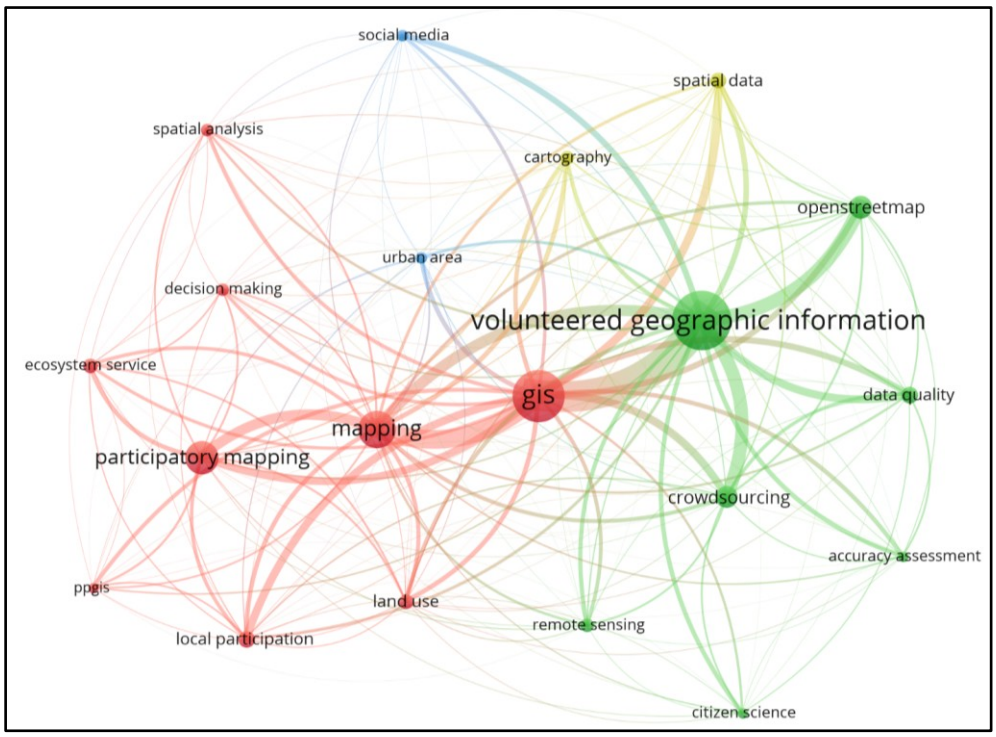


Figure 8: Co-keyword network visualization based on occurrence. Built in VOS Viewer.

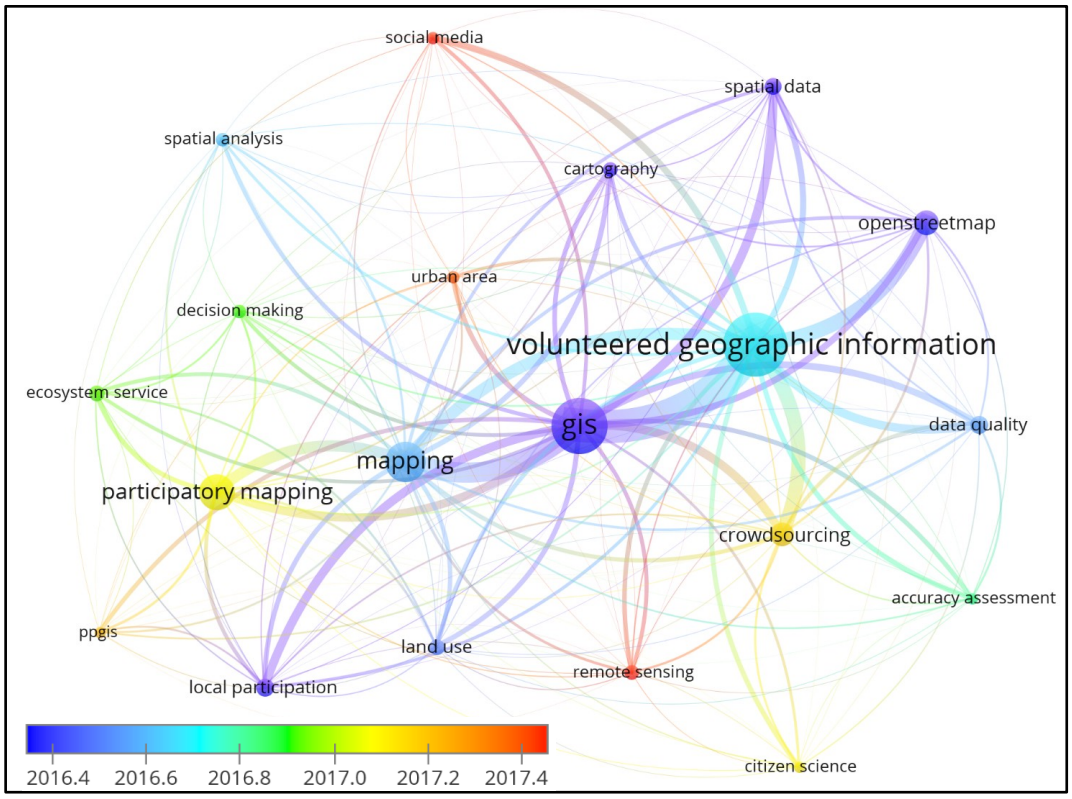


Figure 9: Co-keyword network visualization based on average publication per year. Built in VOS Viewer.

In **Figure 8**, “OpenStreetMap” and “crowdsourcing” are inserted in the same cluster as “Volunteered Geographic Information” and they have a strong link with it. “OpenStreetMap” keyword indicates a main trend among the study field, as it is the currently most used open data platform. Together with “OpenStreetMap” in the green cluster are “accuracy assessment” and “data quality” keywords, indicating the concern with such aspects in the field of VGI. Remote sensing, despite being closer to GIS science, is contained in green cluster, suggesting that some VGI researches are being developed based on remote sensing techniques. The nomenclature of the major area in which VGI is inserted, “citizen science”, also appears in this cluster.

At the same time that direct production or active volunteering is highlighted through “OpenStreetMap” keyword, the terms “social media” and “urban area” inside the same cluster (blue) can represent indirectly production or passive volunteering applied in urban systems management, for example. Yellow cluster, comprising “cartography” and “spatial data” seems to refers to studies related to more technical solutions, migrating from citizen science to geoscience.

In **Figure 9** the colors represents the time varying keyword occurrences. The keywords “social media”, “urban area” and “remote sensing” revealed to be the most actual topics (see average publication per year in **Figure 9**), indicating possible trends among VGI research. “PPGIS” and “Crowdsourcing” are also topics with considerably recent occurrence of average publication per year. On the other hand, “spatial data”, “cartography”, “GIS”, “OpenStreetMap” and “local participation” refers to exhausted keywords, implying more consolidated themes.

As shown in **Figure 8**, after “VGI”, the most occurred keyword is “GIS”, revealing that researchers have focused in VGI more as a Geographic Information Science than as a Citizen Science. However, **Figure 9** demonstrate that “citizen science” has an average publication per year newer than “GIS”, exposing that Citizen Science could be a more up-to-date topic than GIS.

Finally, **Table 4** shows keywords identified in **Figure 8** and their respective occurrence in the 15 most cited papers listed in **Table 3** - contrasted cells indicate that the referred paper contains the keyword among paper (titles, keywords or abstracts).

Table 4: Occurrence of most recurrent keywords among the most cited papers.

	Papers ID (Table 3)														
Keywords	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**	**XI**	**XII**	**XIII**	**XIV**	**XV**
Volunteered Geographic Information		■	■				■		■	■	■	■		■	
Openstreetmap												■		■	
Crowdsourcing		■								■					
Data quality		■												■	
Remote Sensing															
Citizen Science															
Accuracy assessment															
GIS	■			■	■							■			■
Mapping	■		■	■	■	■	■	■	■	■			■		■
Participatory mapping	■			■	■			■					■		■
Ecosystem service	■			■	■			■					■		
Local participation															
Land use					■										
Decision making													■		
Spatial analysis									■						
PPGIS															■
Social media							■		■		■				
Urban area															
Cartography															
Spatial data			■						■	■					

The keywords “remote sensing”, “citizen science”, “accuracy assessment”, “urban area” and “cartography” revealed in the bibliometric map, were not founded among the most cited papers (title, keywords and abstracts). Considering the terms used to compose the initial Scopus search, “volunteered geographic information” was the most recurrent keyword among the most cited papers and most recurrent keywords.

Although “mapping” is substantially noticed in both **Figure 8** and **Table 4**, this can be explained due to the more generic nature of the word and not necessarily to a specific trend or application. The opposite occurs to “participatory mapping” that is focused in a mapping category.

4. Conclusion

This paper presented a comprehensive overview of publications on volunteered geographic information. The 10 years period studied, from 2011 to 2020, returned 1996 publications. From these, 1200 were articles, showing a high level of academic peer reviewed publications related to the topic.

The methodology adopted, based on Scopus Analyse Search Results tool, Bibliometrix and VOS Viewer network construction, can be considered complementary to each other, being improved when used together. Whereas the initial results seem like unrelated tables and diagrams, a critical reading on them showed us curious characteristics. Starting with document types that revealed articles predominance all over the year.

Brandford’s Law was a useful tool in order to identify 16 core sources that comprised 33.33% of all articles published. Another useful parameter was Citescore, allowing to imply a high quality of academic research once, among core sources, 61% of papers were published in journals above Citescore average. VOS Viewer network focusing on countries and authors nationality supports the idea of an intrinsic relationship, mainly when related to the most cited papers. Keywords network was essential in order to understand research field trends and main applications.

The keywords with the strongest links were volunteered geographic information and GIS, implying that besides studying VGI as Citizen Science, researchers have included it in Geographic Information Science wide context. At the same time, Citizen Science, by having the most recent average publication per year, could be moving close to VGI approaches. Some other keywords expose relevant aspects, including trends predicted in literature review: Openstreetmap, data quality, accuracy assessment, social media and crowdsourcing. As also described in introduction section, common applications were confirmed: urban area, land use and ecosystem service.

Although PPGIS was not mentioned in introduction section, together with social media, remote sensing, urban area and crowdsourcing they compose a group of keywords with more recent significant occurrence, indicating trends topics among VGI studies.

Acknowledgments

This study is part of Urban Engineering Postgraduate Program and was financed by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES).

References

BROWN, G. A Review of Sampling Effects and Response Bias in Internet Participatory Mapping (PPGIS/PGIS/VGI). **Transactions in GIS**; v. 21, n. 1, 39-56, 2017. DOI: 21. 10.1111/tgis.12207

BUCZKOWSKI, A. **To pay or not to pay – Google Maps dilemma**, 2012. Available at: <https://geoawesomeness.com/to-pay-or-not-to-pay-google-maps-dilemma/>.

CAPINERI, C.; HAKLAY, M.; HUANG, H.; ANTONIOU, V.; KETTUNEN, J.; FRANK, O.; ROSS, P. Introduction. In: CAPINERI, C.; HAKLAY, M.; HUANG, H.; ANTONIOU, V.; KETTUNEN, J.; FRANK, O.; ROSS, P. (eds.). **European Handbook of Crowdsourced Geographic Information.** London: Ubiquity Press; 2016. p.1–13.

CASTRO, R., TIERRA, A.; LUNA, M. Assessing the horizontal positional accuracy in Openstreetmap: A big data approach. **Advances in Intelligent Systems and Computing**; v. 931, 513–523, 2019. DOI: 10.1007/978-3-030-16184-2_49

CHEN, X.; VO, H.; WANG, Y.; WANG, F. A framework for annotating OpenStreetMap objects using geo-tagged tweets. **GeoInformatica**; v. 22, n. 3, 589–613, 2018. DOI: 10.1007/s10707-018-0323-8

DE ALBUQUERQUE, J.; ECKLE, M.; HERFORT, B.; ZIPF, A. Crowdsourcing geographic information for disaster management and improving urban resilience: an overview of recent developments and lessons learned. In: Capineri, C.; Haklay, M.; Huang, H.; Antoniou, V.; Kettunen, J.; Frank, O.; Ross, P. (eds.). **European Handbook of Crowdsourced Geographic Information.** London: Ubiquity Press; 2016. p.309–321.

DEVKOTA, B.; MIYAZAKI, H.; WITAYANGKURN, A.; KIM, S. Using volunteered geographic information and night-time light remote sensing data to identify tourism areas of interest. **Sustainability (Switzerland)**; v. 11, n. 17, 1-29, 2019. DOI: 10.3390/su11174718.

DORNER, W.; CAMARGO, L. R.; HOFMANN, P. Can geoinformation help to better protect informal settlements? A concept for the city of Medellin. **ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**; v. XLII-3/W8, 115-120, 2019. DOI: 10.5194/isprs-archives-XLII-3-W8-115-2019.

ELSEVIER. Scopus: **The Bibliographic Indexing Leader**, 2021. Available at: <https://www.elsevier.com/research-intelligence/campaigns/get-to-know-scopus/the-bibliographic-indexing-leader>.

ELWOOD, S., GOODCHILD, M. F.; SUI, D. Z. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice, **Annals of the Association of American Geographers**; v. 102, n. 3, 571–590, 2012. DOI: 10.1080/00045608.2011.595657.

FERREIRA, D. Research on big data, VGI, and the tourism and hospitality sector: Concepts, methods, and geographies. In: Sigala, M.; Rahimi, R.; Thelwall, M. (eds.). **Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques and Applications**. Singapore: Springer; 2019. p. 75-85.

FOSSUM, M. **Websites Bypassing Google Maps Due to Fees**, 2012. Available at: <https://www.webpronews.com/websites-bypassing-google-maps-due-to-new-fees>.

GARCIA, X.; GOTTWALD, S.; BENAGES-ALBERT, M.; PAVÓN, D.; RIBAS, A.; VALL-CASAS, P. Evaluating a web-based PPGIS for the rehabilitation of urban riparian corridors. **Applied Geography**; v. 125, 102341, 2020. DOI: 10.1016/j.apgeog.2020.102341

GARFIELD, E. **Citation Indexing**: Its Theory and Application in Science, Technology and Humanities. New York: John Wiley & Sons; 1972.

GOODCHILD, M. F. Citizens as Sensors: The World of Volunteered Geography. **GeoJournal**; v. 69, 211–221, 2007. DOI: 10.1007/s10708-007-9111-y.

HAKLAY, M. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: SUI, D.; ELWOOD, S.; GOODCHILD, M. (eds). **Crowdsourced Geographic Knowledge**. Dordrecht: Springer; 2013. p. 105–122, 2013.

KELER, A.; MAZIMPAKA, J. D. Safety-aware routing for motorised tourists based on open data and VGI. **Journal of Location Based Services**; v. 10, n. (1), 64–77, 2016. DOI: 10.1080/17489725.2016.1170216.

LULEWICZ-SAS, A. Corporate social responsibility in the light of management science - Bibliometric analysis. **Procedia Eng.**; v. 182, 412–417, 2017. DOI: 10.1016/j.proeng.2017.03.124

MENDES, I. A. C.; MARZIALE, M. H. P. O fator de impacto das publicações científica. **Revista Latino-Americana de Enfermagem**; v. 10, n. 4, 466–467, 2002. DOI: 10.1590/S0104-11692002000400001

MIAO, Z.; XIAO, Y.; SHI, W.; HE, Y.; SAMAT, A.; LIU, S. Synergy of Satellite Image and Volunteered Geographic Information to Change Analysis of Impervious Surface. In: International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 10.; 2019, Shanghai, China. **Annals…** Shangai: Hong Kong Polytechnic University, 2019. p. 1-5. DOI: 10.1109/Multi-Temp.2019.8866978

MONDZECH, J.; SESTER, M. Quality Analysis of OpenStreetMap Data Based on Application Needs. **Cartographica: The International Journal for Geographic Information and Geovisualization**; v. 46, n. 2, 115–125, 2011. DOI: 10.3138/carto.46.2.115.

NEIS, P.; ZIELSTRA, D. Generation of a tailored routing network for disabled people based on collaboratively collected geodata; **Applied Geography**, v. 47, 70–77, 2014. DOI: 10.1016/j.apgeog.2013.12.004.

NOVACK, T., WANG, Z.; ZIPF, A. A system for generating customized pleasant pedestrian routes based on openstreetmap data. **Sensors (Switzerland)**; v. 18, n. 11, 3794, 2018. DOI: 10.3390/s18113794.

OKUBO, Y. Bibliometric Indicators and Analysis of Research Systems: Methods and Examples. Paris: OECD Publishing; 1997.

OSM. **OpenStreetMap stats report**, 2020. Available at: <https://www.openstreetmap.org/stats/data_stats.html>.

PANDIAN, Y. **New Wikipedia app for iOS (and an update for our Android App)**, 2012. Available at: <https://diff.wikimedia.org/2012/04/05/new-wikipedia-app-for-ios-and-an-update-for-our-android-app/>.

PAUL, J. D.; HANNAH, D. M.; LIU, W. Editorial: Citizen Science: Reducing Risk and Building Resilience to Natural Hazards. **Frontiers in Earth Science**; v. 7, 320, 2019. DOI: 10.3389/feart.2019.00320.

ROMANELLI, J. P.; FUJIMOTO, J. T.; FERREIRA, M. D.; MILANEZ, H. D. Assessing ecological restoration as a research topic using bibliometric indicators. **Ecological Engineering**; v. 120, 311-320, 2018. DOI: 10.1016/j.ecoleng.2018.06.015.

SANTIN, D. M.; VANZ, S. A. de S.; CAREGNATO, S. E. The analysis of scientific collaboration networks based on bibliometric indicators. In: FRANCO, S. R. K.; FRANCO, M. E. D. P.; LEITE, D. B. C. (eds). **Higher education and knowledge in the centenary of the Córdoba Reform: new perspectives in emerging contexts**. Porto Alegre: EDIPUCRS, 2019. p. 189-207.

SCHMITZ, S.; ZIPF, A.; NEIS, P. **New Applications based on collaborative geodata: the case of Routing**, 2008. Available at: <https://www.semanticscholar.org/paper/New-Applications-based-on-collaborative-geodata-–-Schmitz-Zipf/0cae7b250a72aa477725577ab2badf65d2f94bca>.

SENTURK, I. F.; KEBE, G. Y. A New Approach to Simulating Node Deployment for Smart City Applications Using Geospatial Data. In: International Symposium on Networks, Computers and Communications (ISNCC), 2019, Istanbul, Turkey. **Annals…** Istanbul: IEEE, 2019. p. 1-5. DOI: 10.1109/ISNCC.2019.8909101

VAN ECK, N. J.; WALTMAN, L. Visualizing Bibliometric Networks. In: DING, Y.; ROUSSEAU, R.; WOLFRAM, D. (eds). **Measuring scholarly impact: Methods and practice**. Cham: Springer, 2014. p. 285–320.

WANG, X.; FANG, Z.; SUN, X. Usage patterns of scholarly articles on web of science: A study on web of science usage count. **Scientometrics**; v. 109, 917–926, 2016. DOI: 10.1007/s11192-016-2093-0

ZHAO, D.; STROTMANN, A. Evolution of Research Activities and Intellectual Influences in Information Science 1996–2005: Introducing Author Bibliographic-Coupling Analysis. **Journal of the American Society for Information Science and Tecnhology**, v. 59, 13, 2070-2086, 2008.

ZIPF, A.; MOBASHERI, A.; ROUSELL, A.; HAHMANN, S. Crowdsourcing for individual needs – the case of routing and navigation for mobility-impaired persons. In: CAPINERI, C.; HAKLAY, M.; HUANG,

H.; ANTONIOU, V.; KETTUNEN, J.; FRANK, O.; ROSS, P. (eds.). **European Handbook of Crowdsourced Geographic Information.** London: Ubiquity Press; 2016. p. 325–337.