Supplementary File

Appendix 1. Article collection and filtering.

This work searches for articles published online between 20 November 2007 and 20 November 2017 in 26 journals in the GIScience community. Only the journals included in the Science Citation Index (SCI), Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), and Emerging Sources Citation Index (ESCI) are included. The four indices are among the core collections of Web of Science according to Clarivate Analytics (http://mjl.clarivate.com/). The journals focus more on remote sensing and photogrammetry (e.g., *ISPRS Journal of Photogrammetry and Remote Sensing*) are not included because this review focuses on publications upon GIS.

The keyword ‘VGI’ instead of ‘volunteered geographic information’ is used for the search of articles about volunteered geographic information (VGI). Thus, articles in which the term ‘volunteered geographic information’ is unabbreviated (a low-frequency term) are not included in this review, which are considered of minor relevance to VGI. Other terms related to VGI (e.g., user-generated content, social media, and neogeography) are not used as keywords for the literature search, as this review focuses on VGI as the primary perspective or point of research. The keyword ‘VGI’ serves as a central point (‘seed’) for us to identify other terms (perspectives) related to VGI from articles involving the term ‘VGI’. Except for *GIScience & Remote Sensing* and *Journal of Geographical Systems*, all the remaining 24 journals have returned articles involving the term ‘VGI’. Only research articles and review articles are retrieved. Other contributions such as commentaries, editorials, project reports, or communications are excluded. Articles in which ‘VGI’ is not the abbreviation of ‘volunteered geographic information’ are ignored. This results in 374 articles.

From these 374 articles, we have further manually removed 28 articles in which VGI plays a minor role, e.g., VGI is briefly mentioned in the discussion (Frazier *et al.* (2018) or is a related topic rather than the focus (Brovelli *et al.* (2015). The article filtering process results in 346 articles (326 research articles and 20 review articles) in which VGI is the main topic of exploration or at least is used as a source of data. Each of the 20 review articles is about a sub-topic of VGI, such as VGI quality assessment methods (Senaratne *et al.* 2017) and VGI for natural hazards (Klonner *et al.* 2016); none provides a comprehensive review and is thus treated as the input in the analysis of this review.

Appendix 2. Spatial distribution of the articles related to VGI.

The spatial distribution of the 346 articles is mapped using the institute locations of the first authors of the articles (from 33 countries and 157 research institutes, Figure A1). Most of the articles are published by the researchers with affiliations in the USA or in Europe (Figure A1a). At a finer granularity, most of the articles in Europe are from Germany and the UK (Figure A1b). Table A1 shows the details at individual country and research institute level

Supplementary File

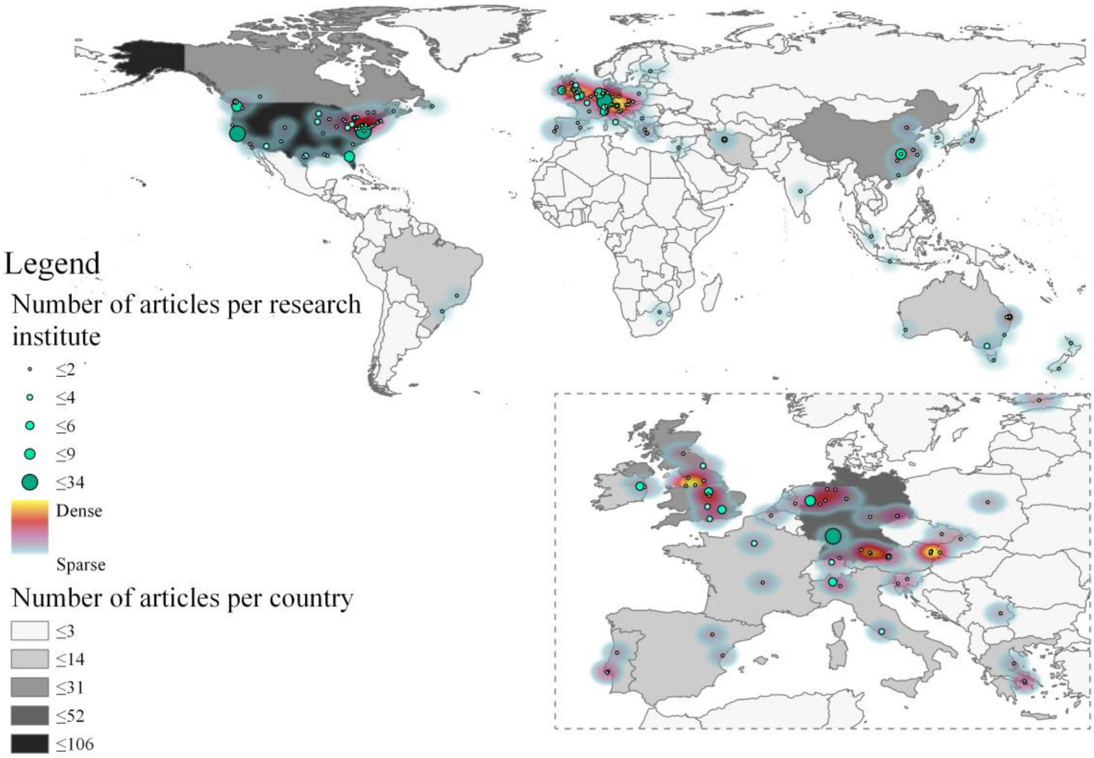

Figure A1. (a) Spatial distribution of the articles across the globe. (b) Inset showing an enlarged view about the articles' distribution in European countries. Heatmap symbology has been enabled for the number of articles per research institute, using ArcGIS Pro 2.1.0® (ESRI Products, Redlands, CA) to show the spatial density.

Table A1. Number of articles (346 articles in total) published by individual countries and research institutes, in descending order.

ID	*Country*	*Number of articles per country*	*Research institute*	*Number of articles per research institute*
1	The USA	105	Heidelberg University, Germany	34
2	Germany	52	George Mason University, the USA	17
3	UK	31	University of California, Santa Barbara, the USA	16
4	China	23	University of Florida, the USA	9
5	Canada	22	University of Twente, the Netherlands	8

6	Australia	14	University of Washington, the USA	8
7	Austria	12	Wuhan University, China	7
8	Italy	11	European Commission Joint Research Centre, Italy	6
9	The Netherlands	9	University College London, the UK	6
10	Ireland	7	National University of Ireland, Maynooth, Ireland	5
11	Spain	7	University of Nottingham, the UK	5
12	Brazil	6	Newcastle University, the UK	4
13	France	6	The French national mapping agency, France	4
14	Iran	5	The Ohio State University, the USA	4
15	Portugal	5	The University of Melbourne, Australia	4
16	Greece	4	University of Oxford, the UK	4
17	Switzerland	4	University of São Paulo, Brazil	4
18	Japan	3	University of Zurich, Switzerland	4
19	Belgium	2	Wilfrid Laurier University, Canada	4
20	New Zealand	2	Arizona State University, the USA	3
21	Poland	2	Iowa State University, the USA	3
22	Slovakia	2	K. N. Toosi University of Technology, Iran	3
23	Singapore	2	Kent State University, the USA	3
24	Czech Republic	1	National Research Council of Italy, Italy	3
25	Finland	1	Simon Fraser University, Canada	3
26	India	1	Texas State University, the USA	3

27	Indonesia	1	University of Melbourne, Australia	3
28	Israel	1	University of Minnesota, the USA	3
29	Mexico	1	University of Portsmouth, the UK	3
30	The Republic of Korea	1	University of Salzburg, Austria	3
31	Serbia	1	AIT Austrian Institute of Technology, Austria	2
32	Slovenia	1	Central South University, China	2
33	South Africa	1	China University of Geosciences, China	2
34			Friedrich-Alexander-University Erlangen-Nürnberg, Germany	2
35			Ghent University, Belgium	2
36			Hofstra University, the USA	2
37			International Institute for Applied Systems Analysis, Austria	2
38			Leibniz Institute of Ecological Urban and Regional Development, Germany	2
39			McGill University, Canada	2
40			Memorial University of Newfoundland, Canada	2
41			Nanjing Normal University, China	2
42			National University of Singapore, Singapore	2
43			Peking University, China	2
44			Politecnico di Milano, Italy	2
45			Salzburg Research Forschungsgesellschaft m.b.H., Austria	2
46			Sun Yat-sen University, China	2

47		The Pennsylvania State University, the USA	2
48		The University of Tokyo, Japan	2
49		Universitat Autònoma de Barcelona, Spain	2
50		Universitat Jaume I, Spain	2
51		University College Dublin, Ireland	2
52		University of British Columbia, Canada	2
53		University of Calgary, Canada	2
54		University of Coimbra, Portugal	2
55		University of Illinois at Urbana–Champaign, the USA	2
56		University of Leeds, the UK	2
57		University of Münster, Germany	2
58		University of Queensland, Australia	2
59		University of Tehran, Iran	2
60		University of Waterloo, Canada	2
61		University of Wisconsin-Madison, the USA	2
62		Vienna University of Technology, Austria	2
63		Zhejiang University, China	2
64		52°North Initiative for Geospatial Open Source Software, Germany	1
65		Adam Mickiewicz University in Poznań, Poland	1
66		Association of American Geographers, the USA	1
67		Beijing Institute of City Planning, China	1

68		Bloomsburg University, the USA	1
69		Bundeswehr University Munich, Germany	1
70		Carleton University, Canada	1
71		Cátedras Conacyt - UNAM, Mexico	1
72		Central Washington University, the USA	1
73		Chinese Academy of Sciences, China	1
74		Claremont Graduate University, the USA	1
75		Clark University, the USA	1
76		Federal University of Parana, Brazil	1
77		Finnish Geospatial Research Institute, National Land Survey of Finland, Finland	1
78		Forest Research Institute, Sekocin Stary, Poland	1
79		Forstburg State University, the USA	1
80		FTW Telecommunications Research Center Vienna, Austria	1
81		Hellenic Military Academy, Greece	1
82		Humboldt State University, the USA	1
83		Institute for Global Environmental Strategies, Japan	1
84		Instituto Politécnico de Coimbra, Portugal	1
85		Lancaster University, the UK	1
86		Leibniz Universität Hannover, Germany	1
87		Liverpool John Moores University, the UK	1
88		Loughborough University, the UK	1

89	Louisiana State University, the USA	1
90	Ludwig-Maximilians-Universität München, Germany	1
91	Massachusetts Institute of Technology, the USA	1
92	Michigan State University, the USA	1
93	Microsoft India Development Center, India	1
94	Nanjing Tech University, China	1
95	National Geospatial Technical Operations Center, the USA	1
96	National Technical University of Athens, Greece	1
97	National University of Defense Technology, China	1
98	Palacký University Olomouc, Czech Republic	1
99	Penn State University, the USA	1
100	Purdue University, the USA	1
101	Queen’s University, Canada	1
102	Queensland University of Technology, Australia	1
103	Rutgers University–New Brunswick, the USA	1
104	Ryerson University, Canada	1
105	San Diego State University, the USA	1
106	Slovak Academy of Sciences, Slovakia	1
107	Surveying and Mapping Authority of the Republic of Slovenia, Slovakia	1

108		Technical University Munich, Germany	1
109		Technical University of Madrid, Spain	1
110		Technion - Israel Institute of Technology, Israel	1
111		Technische Universität Dresden, Germany	1
112		Temple University, the USA	1
113		The State University of New Jersey	1
114		TSCF, France	1
115		Universidade Federal de Minas Gerais, Brazil	1
116		Universidade Nova de Lisboa, Portugal	1
117		Universitas Gadjah Mada, Indonesia	1
118		Université Côte d'Azur, France	1
119		Université Laval, Canada	1
120		University of Auckland, New Zealand	1
121		University of Augsburg, Germany	1
122		University of Bremen, Germany	1
123		University of Colorado, the USA	1
124		University of Colorado at Boulder, the USA	1
125		University of Connecticut, the USA	1
126		University of Edinburgh, the UK	1
127		University of Jaén, Spain	1
128		University of Jena, Germany	1
129		University of Konstanz, Germany	1
130		University of Leicester, the UK	1
131		University of Liverpool, the UK	1

Supplementary File

132		University of Manchester, the UK	1
133		University of Michigan, the USA	1
134		University of New Brunswick, Canada	1
135		University of New Mexico, the USA	1
136		University of New Orleans, the USA	1
137		University of Nis, Serbia	1
138		University of Oldenburg, Germany	1
139		University of Osnabrueck, Germany	1
140		University of Otago, New Zealand	1
141		University of Pretoria, South Africa	1
142		University of Science and Technology, Republic of Korea	1
143		University of South Carolina, the USA	1
144		University of Southern Queensland, Australia	1
145		University of Sydney, Australia	1
146		University of Tasmania, Australia	1
147		University of Texas at San Antonio, the USA	1
148		University of Thessaly, Greece	1
149		University of Trieste, Italy	1
150		University of Western Australia, Australia	1
151		University of Zaragoza, Spain	1
152		University of Žilina, Slovakia	1
153		Wageningen University, the Netherlands	1
154		West Virginia University, the USA	1
155		Ydreams, Portugal	1

156		University of California, Merced, the USA	1
157		University of California, Berkeley, the USA	1

Appendix 3. Latent Dirichlet allocation topic modeling

In this study, the latent Dirichlet allocation (LDA) modeling is conducted using the Gensim Python wrapper for LDA from MALLET (McCallum 2002, Řehůřek 2018). Through LDA, the semantics in the abstracts of the 346 articles are represented by a distribution over topics estimated from the abstracts. Each topic is represented by a distribution over words. Figure A2 shows how the LDA model works using a graphical model. The LDA model distinguishes between similar phrases with different contexts and assigns them to separate topics. Given M documents (i.e., abstracts), for a document d composed of N words (W), a distribution θ is sampled over K topics from a Dirichlet distribution based on a hyper-parameter α, denoted as Dirichlet(α). Thus, the topic distribution for each document d is determined by $\theta_d \sim$ Dirichlet(α). For the K topics, each topic k is associated with a distribution φ over words. The φ is derived from another Dirichlet distribution based on a hyper-parameter β, denoted as Dirichlet(β). Thus, the word distribution of each topic k is determined by $\varphi_k \sim$ Dirichlet(β).

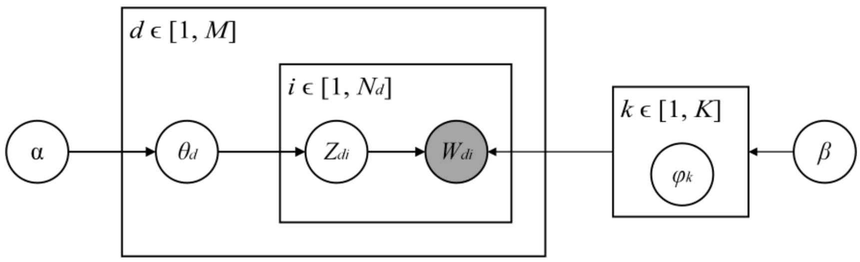

Figure A2. LDA graphical model according to Blei *et al.* (2003). α is a Dirichlet parameter prior on the document topic distribution. β is a Dirichlet parameter prior on the topic word distribution. θ_d denotes the topic distribution for document d. k represents a topic (K topics in total). Z_{di} is the specific topic k associated with an individual word W_{di} in document d. φ_k denotes the word distribution for topic k. M represents the total number of documents (i.e., abstracts). N represents the length of each document.

Pre-processing for LDA topic modeling

To derive the topics meaningfully from the 346 abstracts through LDA topic modeling, following Steiger *et al.* (2016), the texts are pre-processed using natural language processing approaches including tokenization, stop word removal, and lemmatization. These steps reduce the semantic dimension of the raw abstracts, creating word vectors.

Supplementary File

Cohesive strings from the abstracts are split up into single words, or 'tokens', through a tokenization process (Metke-Jimenez *et al.* 2011). Subsequently, the stop words which refer to frequently occurring short-function words without valuable content are removed to reduce noise. The list of stop words from the NLTK natural language toolkit are used (Bird and Loper 2004). Apart from this list of stop words, 'VGI' and 'volunteered', 'geographic', and 'information' are also removed from the token list because our purpose is to derive specific topics under VGI rather than VGI itself. Lastly, lemmatization is performed to convert the words to their root form to simplify further analyses. The lemmatization is conducted using the Python library of spaCY (2018).

Hyper-parameters and number of topics

According to Steyvers and Griffiths (2007), the two hyper-parameters of Dirichlet prior, α and β for this study, are assigned as $50/K$ (K refers to the number of topics) and 0.01, respectively. Gibbs sampling, a form of Markov chain Monte Carlo (MCMC) is used for the LDA posterior parameter inference and optimization. Details about this technique can be found in Griffiths and Steyvers (2004).

One challenge of applying LDA is to determine the number of topics K. A very small K may lead to generic topics that involve many independent yet interrelated small topics. On the contrary, a very large K may either lead to many trivial topics or lead to topics that are difficult to interpret. In this study, the K is determined based on the topic coherence proposed by Röder *et al.* (2015). A topic coherence score distinguishes between good and bad topics, providing a convenient measure to evaluate the performance of a given topic model. A greater topic coherence score indicates a better model performance. In this study, one hundred LDA models are constructed, with the K values ranging from one to 100. The corresponding topic coherence scores are shown in Figure A3. Starting with one topic, they increase rapidly to reach a peak at approximately 0.39 with the topic number K of four. However, four topics would mean big (generic) topics. It is further observed that the coherence scores fluctuate strongly for the K values ranging from four to around 50. From K of 50 onward, the topic coherence scores have stabilized. Therefore, the K is set as 50 for the LDA topic modeling in this study.

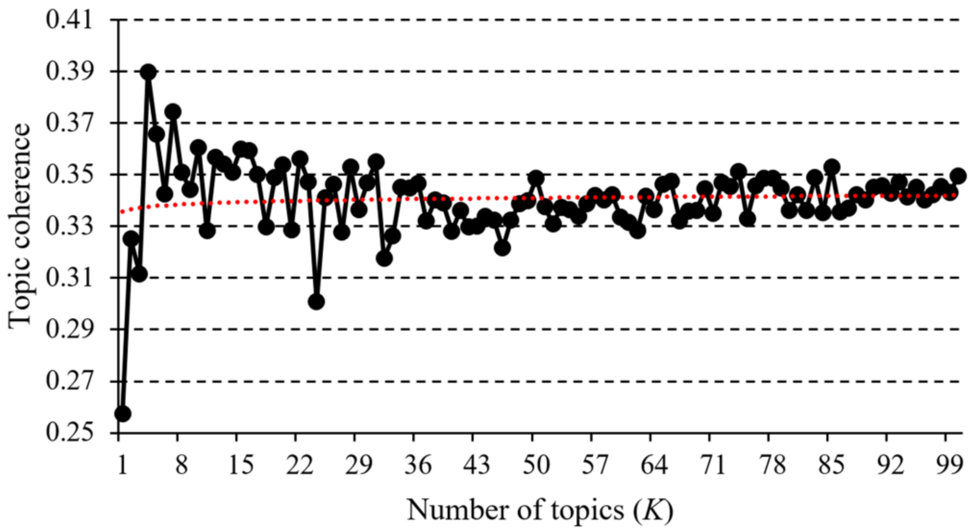

Figure A3. Topic coherence scores of the 100 LDA models with topics numbers (K) ranging from one to 100. Red dash line is a logarithmic trend line.

Supplementary File

Topic labeling, reclassification, and visualization

Each of the 50 topics generated by the LDA topic modeling is represented by a set of keywords (Table A2). The 50 sets of keywords and the respective associated articles are interpreted manually by us to come up with 50 meaningful topic names. The LDA topic modeling reveals the most salient topic of an article and classifies the 346 articles into the 50 topics. Then the LDA article classification results are fine-tuned (readjusted) manually by us according to the topic names and the articles' actual content. Thus, the entire article classification process into the 50 topics is semi-automated, involving both machining learning and human interpretation. Two articles associated with each of the topics are cited in Table A2 for exemplification. The articles associated with the 50 topics are also plotted (Figure A4).

Lastly, based on our domain knowledge on VGI, we notice that many of the 50 topics still have certain commonalities. For example, both Topic 9 ('OSM contribution and contributor patterns') and Topic 31 ('OSM contributor patterns') involve 'OSM contributor patterns'; Topic 49 ('Pest management through smart phone') can be included in Topic 37 ('Environmental monitoring'). Therefore, similar topics are manually grouped together (reclassified), which reduces the number of topics to 13 for a hierarchical topic review.

Supplementary File

Table A2. Fifty topics generated by the LDA topic modelling and the corresponding LDA-generated keywords and number of associated articles. Two articles associated with each of the topics are cited for exemplification. The recommended publication years by the respective journals instead of the first-online publication years are used in the citations.

Topic ID	***Topic label***	***Keywords***	***Number of associated articles***	***Two related articles***
1	OSM data quality	datum, quality, osm, accuracy, study, dataset, area, completeness, base, openstreetmap	25	Koukoletsos *et al.* (2012), Fan *et al.* (2014)
2	Disaster, crisis, emergency, and hazard management	disaster, management, time, datum, crisis, emergency, real, hazard, system, response	25	Poorazizi *et al.* (2015), Chen *et al.* (2016)
3	VGI data quality and credibility	datum, quality, data, generate, model, credibility, increase, research, source, include	17	Bordogna *et al.* (2016a), Koswatte *et al.* (2018)
4	Photograph-based VGI	photo, flickr, photograph, user, poi, panoramio, region, image, geotagg, interest	13	Sun *et al.* (2015a), Alivand and Hochmair (2017)
5	Land use/land cover	land, cover, class, classification, global, use, product, validation, result, map	12	Jokar Arsanjani *et al.* (2013), Schultz *et al.* (2017)
6	Citizen science	citizen, research, science, tool, community, paper, technology, social, open, develop	11	Brandeis and Carrera Zamanillo (2017), Marchante *et al.* (2017)
7	Geography, neogeography, and Web 2.0	web, technology, digital, social, online, neogeography, platform, geography, year, data	11	Haklay (2013), Leszczynski (2014)
8	Conceptual understanding	concept, datum, spatial, user, temporal, mapping, framework, research, provide, influence	10	Rehrl and Gröchenig (2016), Aubrecht *et al.* (2017)
9	OSM contribution and contributor patterns	osm, project, contribution, openstreetmap, contributor, year, area, edit, member, collaborative	10	Neis and Zipf (2012), Jokar Arsanjani and Bakillah (2015)
10	Social media, Twitter	social, medium, study, content, event, characteristic, dimension, form, space, twitter	10	Crooks *et al.* (2012), Stefanidis *et al.* (2013)
11	Urban studies	urban, city, base, network, space, location, area, planning, propose, time	10	Sun *et al.* (2015b), Salvini and Fabrikant (2016)
12	Semantic or geospatial data matching	datum, match, approach, dataset, propose, similarity, integrate, step, article, attribute	9	Mohammadi and Malek (2015), Chehreghan and Ali Abbaspour (2018)
13	Building block level urban studies	building, urban, population, datum, level, scale, city, estimate, coverage, distribution	8	Foster and Dunham (2015), Long *et al.* (2016)
14	Geoweb 2.0 (VGI and Geoweb)	geoweb, web, application, mapping, knowledge, context, digital, government, crowdsourc, paper	8	Johnson and Sieber (2012), Sieber *et al.* (2016)

15	New models, methods, or approaches to handle VGI	datum, spatial, method, approach, pattern, data, study, propose, result, develop	8	Gröchenig *et al.* (2014), Mozas-Calvache (2016)
16	Routing and navigation	study, region, route, pattern, factor, navigation, high, movement, model, task	8	Alivand *et al.* (2015), Jossé *et al.* (2017)
17	System, technology, or data integration	spatial, integration, system, geospatial, research, paper, crowd, community, base, infrastructure	8	Mooney and Corcoran (2014), Zhou *et al.* (2015)
18	User-generated spatial content	user, base, spatial, content, environment, urban, mobile, generate, view, present	8	Khoshamooz and Taleai (2017), Kalvelage *et al.* (2018)
19	VGI, public participation GIS, and participatory GIS	project, participation, participatory, public, practice, study, system, decision, case, process	8	Tulloch (2008), Brown (2017)
20	Digital or data divide	divide, issue, generate, geography, digital, geographical, user, datum, individual, content	7	Cinnamon and Schuurman (2013), Leidig *et al.* (2016)
21	Gazetteer	gazetteer, medium, social, place, source, name, message, build, digital, text	7	Oliveira *et al.* (2016), Gao *et al.* (2017b)
22	Road network	road, network, street, line, segment, match, polygon, rout, pois, graph	7	Yang and Zhang (2015), Tian *et al.* (2018)
23	Social sensing and human activities or behavior	area, identify, sensing, activity, associate, social, image, environment, analysis, sense	7	Sagl *et al.* (2014), Gao *et al.* (2017a)
24	Spatial data infrastructure	datum, geospatial, service, sdi, web, support, infrastructure, development, share, data	7	Budhathoki *et al.* (2008), Bordogna *et al.* (2016b)
25	Accessibility mapping and navigation	accessibility, system, navigation, obstacle, include, make, traditional, provide, individual, present	6	Rice *et al.* (2013), Qin *et al.* (2018)
26	GIScience research agenda with VGI	discussion, phase, science, future, GIScience, geography, research, geographer, discipline, follow	6	Sui (2011), Sui (2014)
27	Local or indigenous knowledge acquisition from VGI	knowledge, field, perspective, framework, spatial, potential, participatory, participant, offer, discuss	6	Quinn and Yapa (2016), Verplanke *et al.* (2016)
28	Data quality measurement or assurance	quality, present, map, datum, set, metric, define, specific, mapping, assure	6	Triglav *et al.* (2011), Yan *et al.* (2017b)
29	Participatory mapping	map, regard, tool, survey, participant, usability, test, category, interface, citizen	6	Aditya (2010), Behrens *et al.* (2015)
30	Public participation GIS, VGI, and sociopolitical geography	participation, public, practice, associate, society, production, geographer, political, literature, geosocial	6	Lin (2013), Cochrane *et al.* (2017)
31	OSM contributor patterns	osm, contributor, feature, object, area, openstreetmap, tag, map, find, place	5	Davidovic *et al.* (2016), Quinn (2016)
32	Sensor network	datum, sensor, make, platform, network, application, develop, article, collect, model	5	De Longueville *et al.* (2010), Regalia *et al.* (2016)

33	VGI user behavior	user, research, study, analysis, result, mapping, potential, focus, base, aim	5	Ooms *et al.* (2016), Poplin *et al.* (2017)
34	Collective map design	map, design, process, cartographic, intelligence, collective, method, clutter, practice, experience	4	Jones *et al.* (2014), Spielman (2014)
35	Surveying and geomatics	map, database, web, domain, reference, aim, survey, internet, effort, geomatic	4	Li *et al.* (2018a), McCartney *et al.* (2015)
36	Collaborative mapping	map, mapping, local, produce, change, focus, geographical, evaluation, future, update	3	Turk (2017), Dalton (2018)
37	Environmental monitoring	datum, project, monitoring, collection, environmental, forest, protocol, monitor, envirocar, purpose	3	Gouveia and Fonseca (2008), Connors *et al.* (2012)
38	Habitat suitability mapping	citizen, tick, map, bite, location, suitability, pattern, habitat, historical, environmental	3	Zhu *et al.* (2015), Zhang *et al.* (2018)
39	Indoor location-based service	service, demand, indoor, system, spatial, public, detect, paper, evacuation, efficiency	3	Goetz and Zipf (2012), Coleman *et al.* (2016)
40	Information communication technologies and social sciences	ict, social, research, exist, result, framework, identify, share, identification, location	3	Janelle (2012), Huggins and Frosina (2017)
41	Spatiotemporal human behavior and demographics monitoring	study, demographic, provide, investigate, online, people, smartphone, dataset, base, represent	3	Chow *et al.* (2012), Li *et al.* (2018b)
42	Collaborative ontology engineering	ontology, concept, methodology, energeer, prototype, framework, knowledge, datum, user, collaborative	2	Janowicz (2012), Kalbasi *et al.* (2014)
43	Earth observation	earth, observation, measurement, vision, digital, device, radiation, scientific, resolution, high	2	Ferster and Coops (2013), Hillen and Höfle (2015)
44	Event mapping	event, web, strategy, volunteer, application, develop, design, tool, demand, framework	2	Panteras *et al.* (2015), Polous *et al.* (2015)
45	Metadata	base, approach, geo, metadata, source, present, web, system, create, set	2	Kalantari *et al.* (2014), Sidda *et al.* (2014)
46	New algorithms of automated map labeling	algorithm, measure, approach, model, low, work, time, label, automate, average	2	Rylov and Reimer (2014), Rylov and Reimer (2017)
47	Privacy concern	Privacy, datum, individual, location, perception, pattern, personal, sensor, model, task	2	Ricker *et al.* (2015), Seidl *et al.* (2016)
48	Map database quality	map, feature, datum, database, source, quality, change, method, spatial, object	1	Maguire and Tomko (2017)
49	Pest management through smart phone	management, ipm, big, mobile, consumer, time, pest, phone, obtain, real	1	Yan *et al.* (2017a)
50	Spatial data fusion	spatial, datum, fusion, formalization, decision, web, discuss, process, pattern, support	1	Wiemann (2017)

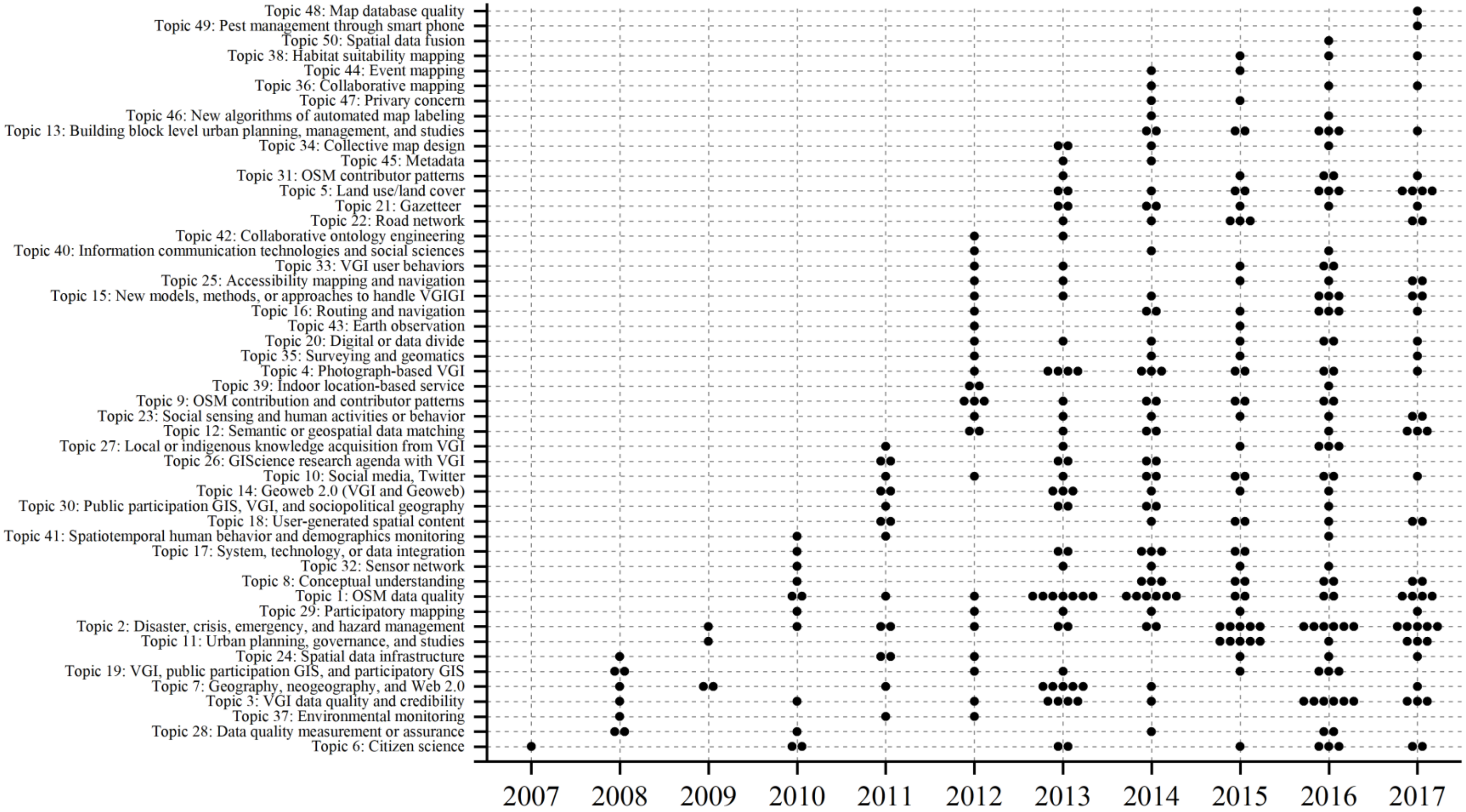

Figure A4. The temporal distribution of the 50 topics over the 10 years of VGI publications (2007-2017). Multiple articles produced in a single year are arranged horizontally, extending from the central cross point (offset plotting), rather than overlapped with each other.

Supplementary File

References for supplementary file

Aditya, T. 2010. Usability issues in applying participatory mapping for neighborhood infrastructure planning. *Transactions in GIS*, 14 (s1), 119-147.

Alivand, M., Hochmair, H. and Srinivasan, S. 2015. Analyzing how travelers choose scenic routes using route choice models. *Computers, environment and urban systems*, 50, 41-52.

Alivand, M. and Hochmair, H. H. 2017. Spatiotemporal analysis of photo contribution patterns to Panoramio and Flickr. *Cartography and Geographic Information Science*, 44 (2), 170-184.

Aubrecht, C., *et al.* 2017. VGDI – advancing the concept: Volunteered geo-dynamic information and its benefits for population dynamics modeling. *Transactions in GIS*, 21 (2), 253-276.

Behrens, J., van Elzakker, C. P. J. M. and Schmidt, M. 2015. Testing the usability of OpenStreetMap's iD tool. *The Cartographic Journal*, 52 (2), 177-184.

Bird, S. and Loper, E. 2004. NLTK: The natural language toolkit. *In: Proceedings of the Association for Computational Linguistics (ACL) 2004 on Interactive Poster and Demonstration Sessions*, Barcelona, Spain: ACL, 31.

Blei, D. M., Ng, A. Y. and Jordan, M. I. 2003. Latent Dirichlet allocation. *Journal of machine Learning research*, 3 (Jan), 993-1022.

Bordogna, G., *et al.* 2016a. On predicting and improving the quality of volunteer geographic information projects. *International Journal of Digital Earth*, 9 (2), 134-155.

Bordogna, G., *et al.* 2016b. A spatial data infrastructure integrating multisource heterogeneous geospatial data and time series: A study case in agriculture. *ISPRS International Journal of Geo-Information*, 5 (5), 73.

Brandeis, M.-N. W. and Carrera Zamanillo, M. I. 2017. Finding meaningful participation in volunteer geographic information and citizen science: A case comparison in environmental application. *Cartography and Geographic Information Science*, 44 (6), 539-550.

Brovelli, M. A., Minghini, M. and Zamboni, G. 2015. Public Participation GIS: A FOSS architecture enabling field-data collection. *International Journal of Digital Earth*, 8 (5), 345-363.

Brown, G. 2017. A review of sampling effects and response bias in Internet participatory mapping (PPGIS/PGIS/VGI). *Transactions in GIS*, 21 (1), 39-56.

Budhathoki, N. R., Bruce, B. and Nedovic-Budic, Z. 2008. Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal*, 72 (3), 149-160.

Chehreghan, A. and Ali Abbaspour, R. 2018. A geometric-based approach for road matching on multi-scale datasets using a genetic algorithm. *Cartography and Geographic Information Science*, 45 (3), 255-269.

Chen, X., *et al.* 2016. Implementing a real-time Twitter-based system for resource dispatch in disaster management. *GeoJournal*, 81 (6), 863-873.

Chow, T. K. E., *et al.* 2012. Using web demographics to model population change of Vietnamese-Americans in Texas between 2000 and 2009. *GeoJournal*, 77 (1), 119-134.

Cinnamon, J. and Schuurman, N. 2013. Confronting the data-divide in a time of spatial turns and volunteered geographic information. *GeoJournal*, 78 (4), 657-674.

Cochrane, L., *et al.* 2017. Searching for social justice in GIScience publications. *Cartography and Geographic Information Science*, 44 (6), 507-520.

Supplementary File

Coleman, D. J., Rajabifard, A. and Kolodziej, K. W. 2016. Expanding the SDI environment: Comparing current spatial data infrastructure with emerging indoor location-based services. *International Journal of Digital Earth,* 9 (6), 629-647.

Connors, J. P., Lei, S. and Kelly, M. 2012. Citizen science in the age of neogeography: Utilizing volunteered geographic information for environmental monitoring. *Annals of the Association of American Geographers,* 102 (6), 1267-1289.

Crooks, A.*, et al.* 2012. #Earthquake: Twitter as a distributed sensor system. *Transactions in GIS,* 17 (1), 124-147.

Dalton, C. M. 2018. Big data from the ground up: Mobile maps and geographic knowledge. *The Professional Geographer,* 70 (1), 157-164.

Davidovic, N.*, et al.* 2016. Tagging in volunteered geographic information: An analysis of tagging practices for cities and urban regions in OpenStreetMap. *ISPRS International Journal of Geo-Information,* 5 (12), 232.

De Longueville, B.*, et al.* 2010. Digital earth's nervous system for crisis events: Real-time sensor web enablement of volunteered geographic information. *International Journal of Digital Earth,* 3 (3), 242-259.

Fan, H.*, et al.* 2014. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science,* 28 (4), 700-719.

Ferster, C. J. and Coops, N. C. 2013. A review of earth observation using mobile personal communication devices. *Computers & Geosciences,* 51, 339-349.

Foster, A. and Dunham, I. M. 2015. Volunteered geographic information, urban forests, & environmental justice. *Computers, environment and urban systems,* 53, 65-75.

Frazier, A. E., Wikle, T. and Kedron, P. 2018. Exploring the anatomy of Geographic Information Systems and Technology (GIS&T) textbooks. *Transactions in GIS,* 22 (1), 165-182.

Gao, S., Janowicz, K. and Couclelis, H. 2017a. Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS,* 21 (3), 446-467.

Gao, S.*, et al.* 2017b. Constructing gazetteers from volunteered Big Geo-Data based on Hadoop. *Computers, environment and urban systems,* 61, 172-186.

Goetz, M. and Zipf, A. 2012. Using crowdsourced geodata for agent-based indoor evacuation simulations. *ISPRS International Journal of Geo-Information,* 1 (2), 186.

Gouveia, C. and Fonseca, A. 2008. New approaches to environmental monitoring: The use of ICT to explore volunteered geographic information. *GeoJournal,* 72 (3), 185-197.

Gröchenig, S., Brunauer, R. and Rehrl, K. 2014. Digging into the history of VGI data-sets: Results from a worldwide study on OpenStreetMap mapping activity. *Journal of Location Based Services,* 8 (3), 198-210.

Griffiths, T. L. and Steyvers, M. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences,* 101 (suppl 1), 5228-5235.

Haklay, M. 2013. Neogeography and the delusion of democratisation. *Environment and Planning A: Economy and Space,* 45 (1), 55-69.

Hillen, F. and Höfle, B. 2015. Geo-reCAPTCHA: Crowdsourcing large amounts of geographic information from earth observation data. *International Journal of Applied Earth Observation and Geoinformation,* 40, 29-38.

Supplementary File

Huggins, C. and Frosina, N. 2017. ICT-driven projects for land governance in Kenya: Disruption and e-government frameworks. *GeoJournal*, 82 (4), 643-663.

Janelle, D. G. 2012. Space-adjusting technologies and the social ecologies of place: Review and research agenda. *International Journal of Geographical Information Science,* 26 (12), 2239-2251.

Janowicz, K. 2012. Observation-driven geo-ontology engineering. *Transactions in GIS,* 16 (3), 351-374.

Johnson, P. A. and Sieber, R. E. 2012. Motivations driving government adoption of the Geoweb. *GeoJournal*, 77 (5), 667-680.

Jokar Arsanjani, J. and Bakillah, M. 2015. Understanding the potential relationship between the socio-economic variables and contributions to OpenStreetMap. *International Journal of Digital Earth*, 8 (11), 861-876.

Jokar Arsanjani, J., *et al.* 2013. Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science,* 27 (12), 2264-2278.

Jones, K., *et al.* 2014. Visualizing perceived spatial data quality of 3D objects within virtual globes. *International Journal of Digital Earth,* 7 (10), 771-788.

Jossé, G., *et al.* 2017. Knowledge extraction from crowdsourced data for the enrichment of road networks. *GeoInformatica,* 21 (4), 763-795.

Kalantari, M., *et al.* 2014. Geospatial Metadata 2.0 – An approach for volunteered geographic information. *Computers, environment and urban systems,* 48, 35-48.

Kalbasi, R., *et al.* 2014. Collaborative ontology development for the Geosciences. *Transactions in GIS,* 18 (6), 834-851.

Kalvelage, K., *et al.* 2018. Assessing the validity of facilitated-volunteered geographic information: Comparisons of expert and novice ratings. *GeoJournal,* 83 (3), 477-488.

Khoshamooz, G. and Taleai, M. 2017. Multi-domain user-generated content based model to enrich road network data for multi-criteria route planning. *Geographical Analysis,* 49 (3), 239-267.

Klonner, C., *et al.* 2016. Volunteered geographic information in natural hazard analysis: A systematic literature review of current approaches with a focus on preparedness and mitigation. *ISPRS International Journal of Geo-Information,* 5 (7), 103.

Koswatte, S., McDougall, K. and Liu, X. 2018. VGI and crowdsourced data credibility analysis using spam email detection techniques. *International Journal of Digital Earth,* 11 (5), 520-532.

Koukoletsos, T., Haklay, M. and Ellul, C. 2012. Assessing data completeness of VGI through an automated matching procedure for linear data. *Transactions in GIS,* 16 (4), 477-498.

Leidig, M., Teeuw, R. M. and Gibson, A. D. 2016. Data poverty: A global evaluation for 2009 to 2013 - implications for sustainable development and disaster risk reduction. *International Journal of Applied Earth Observation and Geoinformation,* 50, 1-9.

Leszczynski, A. 2014. On the neo in neogeography. *Annals of the Association of American Geographers,* 104 (1), 60-79.

Li, D., Shen, X. and Wang, L. 2018a. Connected Geomatics in the big data era. *International Journal of Digital Earth*, 11 (2), 139-153.

Li, M., *et al.* 2018b. Assessing spatiotemporal predictability of LBSN: A case study of three Foursquare datasets. *GeoInformatica,* 22 (3), 541-561.

Supplementary File

Lin, W. 2013. Volunteered geographic information and networked publics? Politics of everyday mapping and spatial narratives. *GeoJournal*, 78 (6), 949-965.

Long, Y., Shen, Y. and Jin, X. 2016. Mapping block-level urban areas for all Chinese cities. *Annals of the American Association of Geographers,* 106 (1), 96-113.

Maguire, S. and Tomko, M. 2017. Ripe for the picking? Dataset maturity assessment based on temporal dynamics of feature definitions. *International Journal of Geographical Information Science,* 31 (7), 1334-1358.

Marchante, H., *et al.* 2017. Using a WebMapping platform to engage volunteers to collect data on invasive plants distribution. *Transactions in GIS*, 21 (2), 238-252.

McCallum, A. K., 2002. *MALLET: A machine learning for language toolkit* [online]. Available from: http://mallet.cs.umass.edu [Accessed 15 August 2018].

McCartney, E. A., *et al.* 2015. Crowdsourcing the national map. *Cartography and Geographic Information Science,* 42 (sup1), 54-57.

Metke-Jimenez, A., Raymond, K. and MacColl, I. 2011. Information extraction from web services : A comparison of tokenisation algorithms. *In: Proceedings of the SKY2011 Workshop: Discovery and Representation of Runnable Knowledge*, Paris, France: Queensland University of Technology.

Mohammadi, N. and Malek, M. 2015. VGI and reference data correspondence based on location-orientation rotary descriptor and segment matching. *Transactions in GIS*, 19 (4), 619-639.

Mooney, P. and Corcoran, P. 2014. Has OpenStreetMap a role in Digital Earth applications? *International Journal of Digital Earth,* 7 (7), 534-553.

Mozas-Calvache, A. T. 2016. Analysis of behaviour of vehicles using VGI data. *International Journal of Geographical Information Science,* 30 (12), 2486-2505.

Neis, P. and Zipf, A. 2012. Analyzing the contributor activity of a volunteered geographic information project — the case of OpenStreetMap. *ISPRS International Journal of Geo-Information,* 1 (2), 146.

Oliveira, M. G. d., *et al.* 2016. Gazetteer enrichment for addressing urban areas: A case study. *Journal of Location Based Services*, 10 (2), 142-159.

Ooms, K., *et al.* 2016. Education in cartography: What is the status of young people's map-reading skills? *Cartography and Geographic Information Science,* 43 (2), 134-153.

Panteras, G., *et al.* 2015. Triangulating social multimedia content for event localization using Flickr and Twitter. *Transactions in GIS,* 19 (5), 694-715.

Polous, K., *et al.* 2015. OpenEventMap: A volunteered location-based service. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 50 (4), 248-258.

Poorazizi, M., Hunter, A. and Steiniger, S. 2015. A volunteered geographic information framework to enable bottom-up disaster management platforms. *ISPRS International Journal of Geo-Information,* 4 (3), 1389.

Poplin, A., Guan, W. and Lewis, B. 2017. Online survey of heterogeneous users and their usage of the interactive mapping platform WorldMap. *The Cartographic Journal*, 54 (3), 214-232.

Qin, H., Curtin, K. M. and Rice, M. T. 2018. Pedestrian network repair with spatial optimization models and geocrowdsourced data. *GeoJournal*, 83 (2), 347-364.

Quinn, S. 2016. A geolinguistic approach for comprehending local influence in OpenStreetMap. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 51 (2), 67-83.

Quinn, S. and Yapa, L. 2016. OpenStreetMap and food security: A case study in the city of Philadelphia. *The Professional Geographer,* 68 (2), 271-280.

Röder, M., Both, A. and Hinneburg, A. 2015. Exploring the space of topic coherence measures. *In: Proceedings of the 8th Association for Computing Machinery (ACM) International Conference on Web Search and Data Mining,* Shanghai, China: ACM, 399-408.

Regalia, B.*, et al.* 2016. Crowdsensing smart ambient environments and services. *Transactions in GIS,* 20 (3), 382-398.

Rehrl, K. and Gröchenig, S. 2016. A framework for data-centric analysis of mapping activity in the context of volunteered geographic information. *ISPRS International Journal of Geo-Information,* 5 (3), 37.

Řehůřek, R., 2018. *gensim topic modelling for humans: Models.ldamallet – Latent Dirichlet Allocation via Mallet* [online]. Available from: https://radimrehurek.com/gensim/models/ldamallet.html#id2 [Accessed 15 August 2018].

Rice, M. T.*, et al.* 2013. Crowdsourcing techniques for augmenting traditional accessibility maps with transitory obstacle information. *Cartography and Geographic Information Science,* 40 (3), 210-219.

Ricker, B., Schuurman, N. and Kessler, F. 2015. Implications of smartphone usage on privacy and spatial cognition: Academic literature and public perceptions. *GeoJournal,* 80 (5), 637-652.

Rylov, M. and Reimer, A. 2017. A practical algorithm for the external annotation of area features. *The Cartographic Journal,* 54 (1), 61-76.

Rylov, M. A. and Reimer, A. W. 2014. A comprehensive multi-criteria model for high cartographic quality point-feature label placement. *Cartographica: The International Journal for Geographic Information and Geovisualization,* 49 (1), 52-68.

Sagl, G., Delmelle, E. and Delmelle, E. 2014. Mapping collective human activity in an urban environment based on mobile phone data. *Cartography and Geographic Information Science,* 41 (3), 272-285.

Salvini, M. M. and Fabrikant, S. I. 2016. Spatialization of user-generated content to uncover the multirelational world city network. *Environment and Planning B: Planning and Design,* 43 (1), 228-248.

Schultz, M.*, et al.* 2017. Open land cover from OpenStreetMap and remote sensing. *International Journal of Applied Earth Observation and Geoinformation,* 63, 206-213.

Seidl, D. E., Jankowski, P. and Tsou, M.-H. 2016. Privacy and spatial pattern preservation in masked GPS trajectory data. *International Journal of Geographical Information Science,* 30 (4), 785-800.

Senaratne, H.*, et al.* 2017. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science,* 31 (1), 139-167.

Sidda, N. K.*, et al.* 2014. Expeditious management plan towards digital earth. *International Journal of Digital Earth,* 7 (8), 635-649.

Sieber, R. E.*, et al.* 2016. Doing public participation on the geospatial web. *Annals of the American Association of Geographers,* 106 (5), 1030-1046.

spaCY, 2018. *Industrial-strength natural language processing in Python* [online]. Available from: https://spacy.io/ [Accessed 23 August 2018].

Supplementary File

Spielman, S. E. 2014. Spatial collective intelligence? Credibility, accuracy, and volunteered geographic information. *Cartography and Geographic Information Science,* 41 (2), 115-124.

Stefanidis, A., Crooks, A. and Radzikowski, J. 2013. Harvesting ambient geospatial information from social media feeds. *GeoJournal,* 78 (2), 319-338.

Steiger, E., Resch, B. and Zipf, A. 2016. Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. *International Journal of Geographical Information Science,* 30 (9), 1694-1716.

Steyvers, M. and Griffiths, T. 2007. Probabilistic topic models. *Handbook of latent semantic analysis,* 427 (7), 424-440.

Sui, D. 2014. Opportunities and impediments for open GIS. *Transactions in GIS,* 18 (1), 1-24.

Sui, D. Z. 2011. Introduction: Strategic directions for the geographical sciences in the next decade. *The Professional Geographer,* 63 (3), 305-309.

Sun, Y., *et al.* 2015a. Road-based travel recommendation using geo-tagged images. *Computers, environment and urban systems,* 53, 110-122.

Sun, Y., *et al.* 2015b. Identifying the city center using human travel flows generated from location-based social networking data. *Environment and Planning B: Planning and Design,* 43 (3), 480-498.

Tian, J., *et al.* 2018. On the degree correlation of urban road networks. *Transactions in GIS,* 22 (1), 119-148.

Triglav, J., Petrovič, D. and Stopar, B. 2011. Spatio-temporal evaluation matrices for geospatial data. *International Journal of Applied Earth Observation and Geoinformation,* 13 (1), 100-109.

Tulloch, D. L. 2008. Is VGI participation? From vernal pools to video games. *GeoJournal,* 72 (3), 161-171.

Turk, C. 2017. Cartographica incognita: ‘Dijital Jedis’, satellite salvation and the mysteries of the ‘Missing Maps’. *The Cartographic Journal,* 54 (1), 14-23.

Verplanke, J., *et al.* 2016. A shared perspective for PGIS and VGI. *The Cartographic Journal,* 53 (4), 308-317.

Wiemann, S. 2017. Formalization and web-based implementation of spatial data fusion. *Computers & Geosciences,* 99, 107-115.

Yan, Y., Feng, C.-C. and Chang, K. 2017a. Towards enhancing integrated pest management based on volunteered geographic information. *ISPRS International Journal of Geo-Information,* 6 (7), 224.

Yan, Y., Feng, C.-C. and Wang, Y.-C. 2017b. Utilizing fuzzy set theory to assure the quality of volunteered geographic information. *GeoJournal,* 82 (3), 517-532.

Yang, B. and Zhang, Y. 2015. Pattern-mining approach for conflating crowdsourcing road networks with POIs. *International Journal of Geographical Information Science,* 29 (5), 786-805.

Zhang, G., *et al.* 2018. Validity of historical volunteered geographic information: Evaluating citizen data for mapping historical geographic phenomena. *Transactions in GIS,* 22 (1), 149-164.

Zhou, X., *et al.* 2015. Dynamically integrating OSM data into a borderland database. *ISPRS International Journal of Geo-Information,* 4 (3), 1707.

Zhu, A. X., *et al.* 2015. A citizen data-based approach to predictive mapping of spatial variation of natural phenomena. *International Journal of Geographical Information Science,* 29 (10), 1864-1886.