

Volunteered geographic information and crowdsourcing in geographic practice: A unifying conceptual framework

Guiming Zhang

Department of Geography & the Environment, University of Denver, Denver, CO,
United States

27.1 Volunteered geographic information and crowdsourcing in geographic practice

Volunteered geographic information (VGI) broadly refers to geographic data and information contributed by the general public, most of whom are not necessarily trained professionals with expertise in spatial data collection (Zhang, 2021). With the term VGI formally coined by Professor Goodchild (2007), VGI as a research field within GIScience has been in steady growth and proliferation. VGI has become an important source of geospatial (big) data depicting the physical and built environments (and humans therein) to support diverse geographic research and applications (Huang et al., 2024; Yan et al., 2020). Prominent examples include OpenStreetMap, which produces and maintains a freely accessible geographic database for the entire globe (Haklay & Weber, 2008), and iNaturalist, which aims to compile a biodiversity database for species of any taxon on Earth (Unger et al., 2021). Apart from VGI's practical significance in terms of producing geographic data, VGI serves also as a mechanism and propeller to form new social practices that democratize the knowledge production processes because the participation of VGI contributors from different geographies and backgrounds may bring better engagement and representation of social groups and populations who are traditionally underrepresented in such processes (Elwood et al., 2012; Fast, 2024; Sui et al., 2013). This chapter does not intend to provide yet another systematic review of VGI research topics or VGI applications, as readers may find excellent summaries in this regard in existing reviews (Cui et al., 2021; Feng et al., 2022; Fast, 2024; Huang et al., 2024; Senaratne et al., 2017; Yan et al., 2020; Zhang & Zhu, 2018).

In the literature across domains, VGI has many variant names that convey nuanced specificity in particular contexts of the “grass-rooted” geographic data creation and applications (e.g., user-generated content, location-based social media data, geographic citizen science, crowdsourcing geographic knowledge, nonauthoritative data). The variety of the synonyms of VGI certainly represents the broadness and inclusiveness of VGI as an umbrella term and the richness of VGI-related research and applications. However, as argued by [Fast \(2024\)](#), it also reflects a disjointed research field. It may lead to VGI as a research domain to fragmentize or even extinguish, because scholars may be so zoomed in and focused on specific types or instances of VGI (e.g., social media VGI) that they lose interest, intent, and motivation to generalize the investigations in the specialty area to advance our common understanding of the VGI domain at large. Joining ([Fast, 2024](#)), the author also advocates using the term VGI as the unifying language to integrate VGI-related research carried out under the many *dialect* names of VGI, aiming to sustain the coherence of VGI research and facilitate our collective understanding of VGI. To achieve such a goal, this chapter proposes a conceptual VGI research framework to unify VGI-based research and applications.

Considering a particular VGI data record, no matter which type of VGI it falls in, or if it was contributed voluntarily (actively) or unknowingly (passively) by the data contributor, the VGI record likely contains the following components of information ([Zhang & Zhu, 2018](#)). First is the thematic information about a target that was observed and reported. Here, the target can be a feature or phenomenon in the physical and built environments. For example, a VGI record can be about encountering wildlife or sighting a traffic accident. As for media, text, photo, audio, and video could all be used in a VGI observation to convey the thematic information. Second, there is the geographic location and temporal timestamp of the observation. Location and time can be either estimated and described by the human observer or automatically recorded with available enabling technologies such as a GPS-equipped smartphone. Thirdly, there may be information about the observer (self-assessed level of expertise, demographic and socioeconomic status, etc.), if the observer chooses to disclose such information. Lastly, although it is rarely the case, information about the data collection process may be present or implied in the VGI record. In some cases, such information needs to come directly from the contributor. For instance, eBird asks if the submitting birder has reported all the birds seen in a single sampling event ([Sullivan et al., 2009](#)). In other cases, such information is implicit or can be derived. For example, iNaturalist’s published identification of a species observation is based on the discussion and vetting/voting process among community members ([Unger et al., 2021](#)).

One may argue that the four components of VGI are no different from those for traditional (authoritative) geographic data. The thematic information, location, and time components are indeed expected for geographic data, and so is information about the data collector and the data collection process. However, it is the data collector and data collection process that distinguishes VGI from traditional geographic data collection practices. While professionals with qualifications collect authoritative geographic data following standard protocols, anyone in the general public can contribute to VGI, and often, there are no standard protocols governing the data collection process at all. From the perspective of authoritative geographic data, the most important for assessing its data quality and fitness of use is the thematic information it captures and the spatial and temporal references; The underlying data collector and data collection process are unimportant or even invisible because qualified professionals and standard protocols are deemed trustworthy and homogeneous by default. For VGI, though, both the data contributor (volunteer) and data collection process can be very heterogeneous, rendering them

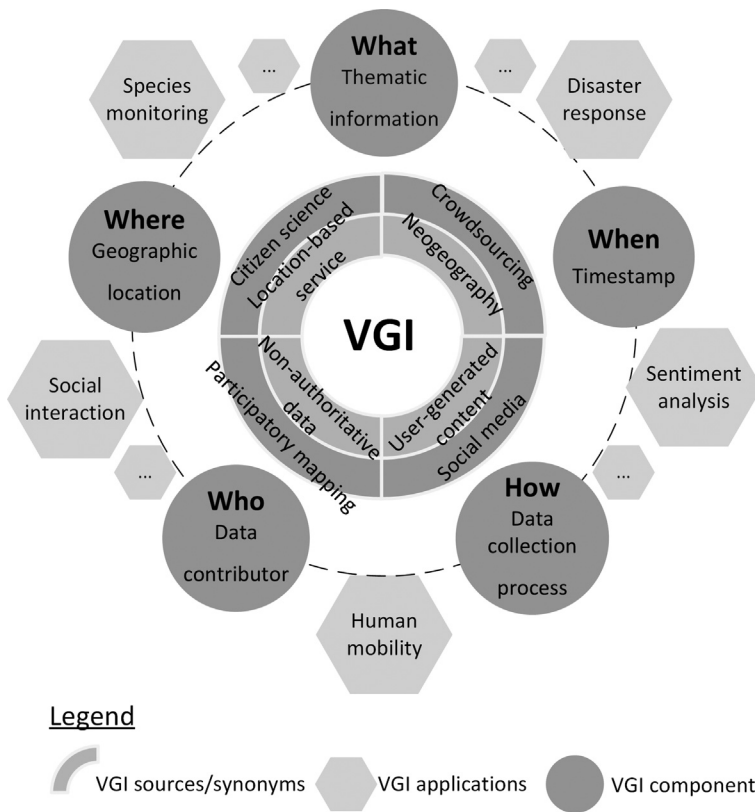


FIGURE 27.1 A diagram illustrating the conceptual VGI research framework. VGI; Volunteered geographic information.

nonnegligible, even prominent, components in VGI data quality assessment and evaluating the fitness of use of VGI data. Lastly, what may be the most striking difference between VGI and traditional geographic data is that VGI data contributors themselves, as a collaborative community, are often the subject of study. Examples are studies on human mobility patterns using social media VGI (Huang & Wong, 2015; Huang et al., 2020) and those examining the patterns of social interactions among VGI contributors (Mooney & Corcoran, 2014; Zhang et al., 2024a).

The five components of VGI form a common thread that can be used to stitch together research and applications involving various types of VGI. The remainder of this chapter presents a conceptual VGI research framework constructed based on these five components of VGI as an attempt to unify existing and future VGI research efforts to enable exploring the potentials and limitations of VGI as an integral research domain within GIScience.

27.2 A unifying conceptual framework for volunteered geographic information research

The conceptual framework for unifying VGI research (Fig. 27.1) uses VGI as an overarching term to encompass various types, sources, and synonyms of VGI, as long as the data is

contributed, either actively or passively, by volunteers who are not contributing data in any form of official capacity or expecting monetary benefits in return (Goodchild, 2007). The conceptual framework, in essence, consists of a series of questions that should be asked regarding the five components of VGI, which are in this chapter referred to as the *what*, *where*, *when*, *who*, and *how* questions. Answering these questions sheds light on assessing VGI data quality as well as evaluating the fitness of the use of VGI for particular applications. These questions are presented in detail in the following section, with a focus on the nature of each question and the implications on VGI data quality and on the fitness of use of VGI for domain applications.

27.2.1 What

The *what* question concerns the thematic information embedded in VGI data (i.e., what is the geographic feature or phenomenon depicted in the VGI data?). The answer to this question largely determines how the VGI data can be useful for various domain applications. For instance, VGI observations of species occurrences are utilized to model and map species distribution over space and time (Fink et al., 2020). VGI records on real-time particulate matter concentration are used to monitor air quality and air pollution (Lu et al., 2022). Sometimes the thematic information can be about people instead of geographic features or phenomena. This is especially true for social media VGI (Feng et al., 2022). Users post on social media to express their opinions and emotions, and these posts can be harnessed to study social media users' attitudes toward certain social events, such as political elections (Jaidka et al., 2019). Users post on social media to seek help in emergency cases such as natural disasters, and such posts can be used to inform disaster responses (Li et al., 2019). Another interesting example is found in (Alizadeh & Behzadan, 2023), where urban street-level flood depth is inferred from crowdsourced street photos of traffic signs by comparing the visible part of stop sign poles before and after the flood event.

The answer to the *what* question in many cases is obvious when VGI projects or platforms are set up with clear objectives. For example, OpenStreetMap is for volunteer mappers to digitize geographic features such as rivers, roads, buildings, etc., from remote sensing imagery (Haklay & Weber, 2008), while eBird is for bird watchers to share their birding records (Sullivan et al., 2009). However, in some other cases, it takes effort to mine the thematic information of interest from VGI records. Take social media, for example, social media platforms are for users to post any sort of opinion and information. If one wants to make use of social media posts to help with disaster damage assessment, one will have to sift through the posts and identify only those texts and images relevant to disaster damage, using natural language processing and computer vision techniques (Feng et al., 2022).

The *what* question can also be asked collectively upon all records collected by a VGI project to profile the thematic composition of the data and reveal potential thematic bias therein. Individual VGI contributor's voluntary observation effort is largely driven by self-interest and preferences, and they decide what to observe at their own discretion. Across biodiversity VGI platforms, observations of common species are of disproportionately higher frequencies in the database compared to rare species because common species are more likely to be encountered by observers (Fitzpatrick et al., 2009). As an extreme example, only approximately 0.01% of iNaturalist observations are on bacteria (Zhang et al., 2024a), while according to a 2017 study, 70%–90% of Earth's species are bacteria (Larsen et al., 2017). One must bear in mind such potential bias when evaluating the fitness of use of VGI data for a particular application.

27.2.2 Where

The *where* question is about the geographic reference of VGI (i.e., where was the VGI observation conducted?). Geographic references are often manifested as point, linear, or areal locations with varied levels of detail. In many cases, when a GPS-equipped smartphone is used in collecting VGI data, the locations are recorded as point geographic coordinates. For instance, photos of species sightings were taken and submitted through the iNaturalist mobile app. Without access to precise geolocating tools, contributors may resort to verbal or textual descriptions to express location reference. For example, a driver may report a car crash on a highway by describing the crash location as between two consecutive (numbered) exits, implying a vague linear location reference. In other cases, VGI records can only be tied to areal locations for reasons such as lacking access to precise geolocating tools or protecting geoprivacy. For instance, when submitting historical birding records to eBird, only geo-referencing to townships or cities is available. Social media users have control over the location granularity of the geotag of their posts to not show the exact point location, but instead, only the neighborhood or city in which they are posting.

Obviously, the level of detail in the geo-reference of VGI implies varied levels of positional accuracy, an important consideration in assessing VGI data quality. As a result, it will also affect the fitness for use of the VGI data for applications. While for certain applications, knowing where things are exactly is necessary to make VGI data useful (e.g., exact location of a crime event for police dispatch), knowing only the general location (linear or areal location references) is sufficient for effectively using VGI in other situations (e.g., bypassing a traffic-jammed highway segment, or counting the total number of species in a township) (Zhu et al., 2015). Moreover, the location reference of VGI is more than representing where things are; it has further implications. When VGI is used along with other spatial data to perform spatial analysis and modeling tasks, VGI locations are geographically overlaid on top of other thematic layers (e.g., land cover map) to answer inquiries involving what is at the locations (e.g., profiling land cover type composition across a wildlife species' occurrence locations).

Additionally, there may be a need to distinguish between where the observer stands and where the feature or phenomenon under observation is located at the time of observation. If the location of a VGI record is automatically recorded using a piece of equipment such as a smartphone or GPS receiver, the recorded location is where the observer stands to observe, which could be different from the actual location of the feature under observation. The difference between the two locations can be significant enough for some applications that it warrants special treatments to "recover" the feature or phenomenon location from the recorded observer location to reduce positional errors. For instance, VGI contributors can only observe wildlife at a distance (e.g., hundreds of meters away). When such VGI-sourced observation locations are used for species distribution modeling to extract environmental variable values from high-resolution environmental covariates (e.g., 30-m raster layers), using observer location in place of actual wildlife location would introduce error into subsequent modeling (Zhang et al., 2018). As another example, in disaster damage assessment, when relying on photos of distant objects (e.g., buildings) taken by citizens with their smartphones, both distance and orientation between the observer and the objects need to be accounted for to infer the true location of the object (Lee et al., 2022).

The *where* question, similar to the *what* question, can also be asked on all records of a VGI dataset to examine the collective spatial pattern of the observations and reveal potential bias. It

is generally true that VGI contributors tend to conduct more observations in areas that are more accessible, which are usually places within the vicinity of large population centers. For example, the sampling efforts of eBird contributors in the United States are very much concentrated in and around large cities (Zhang, 2020). Due to such spatial bias in VGI contributors' observation efforts, the spatial pattern of VGI observations on the target feature or phenomenon does not necessarily reflect its true geographic distribution (Zhang & Zhu, 2018). In VGI applications, measures should be taken to account for such spatial bias to improve the quality of inferences made out of sample data sourced from VGI observations (Zhang & Zhu, 2019a, 2019b).

27.2.3 When

The *when* question has to do with the timestamp of a VGI record (i.e., when was the VGI record created in the database? and when was the VGI observation conducted?). In the following discussion, it is assumed that the timestamp refers to observation time and thus a VGI record reflects the state of the target feature or phenomenon at the time of observation, although the two times may be different and the difference cannot be ignored in some cases (e.g., historical birding records submitted to eBird). The timestamp of VGI observations makes it feasible to study temporal changes of the target feature or phenomenon under observation, for example, how users' level of activeness on social media varies over the hours of a day (Li et al., 2013), how people's interest in a disastrous event expressed on social media change over the weeks following the aftermath of a major hurricane (Kottwitz et al., 2023), and how individual VGI contributor's data contribution behavior may evolve (Zhang, 2024; Zhang et al., 2024a) and be impacted by disruptive events (Ahmouda et al., 2018; Zhang, 2024).

One should bear in mind that VGI observations collectively may not be representative as a sample of the temporal characteristics of the target feature or phenomenon. In other words, temporal bias may exist in VGI observations. VGI contributors carry out observations whenever they want, and it is convenient to them, not necessarily considering the temporal characteristics of the target feature or phenomenon under observation. For instance, it is observed in eBird that significant inter- as well as intra-annual (seasonal) variations exist in the number of birding events, and bird watchers submit significantly more data on weekends compared to weekdays (Zhang, 2020). Such artificial temporal variations in VGI contributors' observation efforts introduce temporal bias in VGI data, which should be assessed and accounted for so meaningful temporal patterns regarding the target feature or phenomenon can be derived.

27.2.4 Who

The *who* question focuses on VGI contributors (i.e., who are the volunteers that actively or passively contribute to VGI?). Generally, each VGI project or platform has a specific theme of the data it intends to collect and therefore attracts contributions from a specific niche of interested volunteers who are willing to engage in the project (social media is an exception, as users can post anything without a specific theme). Amongst the five components of VGI, the who component is arguably the most essential one because it is the who (i.e., VGI contributor) that determines what to observe, where and when to observe, and how to observe (Zhang, 2024).

Knowing who the VGI contributors are is important for at least two reasons. Firstly, it helps end users evaluate VGI data quality and understand where potential data quality issues may come from. VGI contributors are often self-driven individuals who engage with the VGI project

out of their interest in the theme and with a spirit of altruism, assuming that the data they contribute would benefit others. They do not always possess the necessary data collection skills as trained professionals would. As a result, the collective body of VGI contributors is heterogeneous in their data collection expertise and behavior, which renders the resulting VGI data prone to inconsistencies, errors, and uncertainties (Goodchild & Li, 2012). Knowing the VGI contributors, such as their individual expertise level, activity pattern, and taxonomic specialization, would facilitate proper evaluations of VGI data quality (Di Cecco et al., 2021; Haklay et al., 2010; Kendall et al., 1996; Neis & Zipf, 2012; Sauer et al., 1994).

Secondly, VGI contributors themselves could be the subjects of study in VGI applications, where it is the emotions, opinions, or mobilities of VGI contributors that are of interest. For example, social media VGI has long been utilized to examine human mobility (Huang et al., 2020; Jurdak et al., 2015; Liu et al., 2019) and residents' attitudes and perceptions toward place (Cui et al., 2021; Kruse et al., 2021) and even predict the outcomes of political elections (Jaidka et al., 2019). Such VGI applications often call for assessing to what extent the group of people represented in the VGI data (e.g., social media users), as a sample, is representative of the underlying population, so that any pattern extracted from the sample data is generalizable to the general population. In general, VGI contributors are rarely representative of the general population; There are always biases in VGI contributors to fully reflecting the demographic and socioeconomic characteristics of the general population (Huang & Wong, 2015; Li et al., 2013; Malik et al., 2015).

Furthermore, VGI contributors often form collaborative communities and social networks to facilitate VGI data production. In such cases, the social interactions in VGI communities are of interest for investigation because the interactions are considered critical for maintaining and improving the accuracy of the resulting VGI data (Goodchild & Li, 2012; Haklay et al., 2010; McGough et al., 2022). Examples are the inter-contributor co-editing mapping activities in OpenStreetMap and species identification activities in iNaturalist. Understanding the patterns of the interactions and potential drivers sheds light on how the interactions in the respective communities help more accurately map geographic features and identify species records, respectively (Mooney & Corcoran, 2014; Zhang et al., 2024a).

27.2.5 How

The *how* question deals with VGI data collection processes (i.e., how is the VGI data collected?). In a broad sense, the answers to the *what*, *where*, *when*, and *who* questions, as discussed earlier can all be part of the answer to the *how* question. After all, the data collection processes underlying VGI data inevitably involve these four components that drive and shape the processes. Nonetheless, answers to the four questions may not give a full picture of the VGI data collection processes, for which additional questions could always be asked to shed light on. The necessity of asking these questions arises from the fact that VGI data collection efforts often do not follow established standard protocols, and therefore, individual VGI contributors have much flexibility and control regarding how to collect VGI data. Answering these additional questions helps VGI end users better understand how VGI data collection processes unfold at both the individual data contributor and collective community levels (Zhang, 2024).

Taking iNaturalist as an example, many more questions can be asked about a particular species record submitted to iNaturalist. Regarding geographic and temporal references (*where* and *when*), was the geographic location of the record automatically extracted from the photo

taken with a smartphone? Or was it specified by the observer on a map? Was the location obscured to preserve geoprivacy? Similarly, was the timestamp extracted from the photo or hand-typed in? Concerning thematic information (*what*), where did the species identification originate? Was it recommended by the pretrained computer vision model embedded in the iNaturalist mobile app? If this were the case, what was the model's confidence in the identification? With regard to VGI contributors (*who*), is the identification assigned solely by the observer without relying on the computer vision model? What was the profile of the observer like? Does the observer seem to be an expert in observing and identifying species of this taxon (and so assumably implies high confidence in the identification)? Or was the identification proposed by another iNaturalist community member? Does the identifier seem to possess expertise in identifying this species? Furthermore, were there multiple consistent (or contradicting) identifications proposed by community members? Was the final identification vetted or voted on by the community?

The list of questions like these can grow as needed, so VGI data end users can gain pieces of relevant information revealing the processes involved in VGI data production. Eventually, a clear overall picture depicting the underlying data collection processes could emerge, useful for VGI data end users to assess VGI data quality and evaluate the fitness of use of VGI data for certain applications.

27.3 Summary and prospect

Since the term VGI was coined about two decades ago, VGI has been established as a research field within GIScience with a wide variety of applications seen in domains beyond geospatial sciences. As a result, VGI has many *dialect* names (or synonyms) in the literature, depending on the application domain and the need to express specificities in particular contexts of VGI data creation and utilization. The dialect names of VGI, while reflecting the broadness and variety of VGI research and applications, risk fragmentizing VGI as an integral research domain. This chapter presents a conceptual framework as an attempt to unify VGI-related research. Central to this conceptual framework is the use of the term VGI as the unifying *language* to encompass the various types, sources, dialect names, and synonyms of VGI-style geographic data. The framework consists of a series of questions that can and should be asked on the five generic components of VGI data, which are the *what* question on thematic information embedded in VGI, the *where* question on geographic location of VGI observation, the *when* question on time of observation, the *who* question on data contributor, and the *how* question on data collection processes. Investigating these questions helps end users evaluate VGI data quality, understand potential causes of any data quality issues, and ultimately facilitate assessing the fitness of the use of VGI data for specific application purposes. The author argues that research efforts and applications involving any type, source, dialect, or synonym of VGI can be (re)examined through the lenses of this conceptual framework. This conceptual framework, therefore, serves as a universal foundation framework for integrating VGI-related research endeavors and applications to allow exploring the collective potentials and limitations of VGI as an integral research domain.

Looking forward, VGI will undoubtedly keep growing as a major source of geospatial big data. Empowered by technological advancements and social transformations, the general public's capability and awareness of participating in geospatial data creation and geographic

knowledge production are ever-increasing (Sui et al., 2013). There will be an influx of data, information, and knowledge provisioned through VGI, which are useful for better understanding various elements of the physical and built environments (e.g., land cover and land use, biodiversity, air quality, water quality, and traffic conditions) (Huang et al., 2024). Along this line, in data-scarce regions of the world, VGI could serve as an effective mechanism for collecting data to measure and monitor the progress toward the United Nations' sustainable development goals (SDGs) (Solís et al., 2018). Also noticeable in recent years, geospatial artificial intelligence (GeoAI) has emerged as the state-of-the-art method for analyzing and modeling geospatial data to study geographic and social phenomena (Hu et al., 2024). GeoAI often requires huge amounts of data for model training, and VGI could be a perfect "fuel" to provision large-scale yet low-cost training data.

Most VGI projects and platforms so far have been fixated on observing the environments external to human beings, and not enough attention was paid to the volunteers internal to VGI itself. Volunteers are indeed human sensors (Goodchild, 2007) who, in many cases, are actively observing and reporting data on the outside world. Nevertheless, as discussed in Section 27.2.4, they are also self-aware human beings who play very important roles in VGI data creation processes. They often interconnect to form social networks wherein the social interaction among volunteers is an intrinsic part of VGI data production and a necessity to achieve VGI projects' long-term sustainability (Zhang et al., 2024a; Zhang et al., 2024b). The author advocates that more future research efforts should be devoted to investigating the dynamics among the volunteers behind VGI.

References

- Ahmouda, A., Hochmair, H. H., & Cvetojevic, S. (2018). Analyzing the effect of earthquakes on OpenStreetMap contribution patterns and tweeting activities. *Geo-Spatial Information Science*, 21(3), 195–212. <http://www.tandfonline.com/loi/tgsi20>. doi:10.1080/10095020.2018.1498666.
- Alizadeh, B., & Behzadan, A. H. (2023). Scalable flood inundation mapping using deep convolutional networks and traffic signage. *Computational Urban Science*, 3(1). <https://link.springer.com/journal/43762>. doi:10.1007/s43762-023-00090-1.
- Cui, N., Malleson, N., Houlden, V., & Comber, A. (2021). Using VGI and social media data to understand urban green space: A narrative literature review. *ISPRS International Journal of Geo-Information*, 10(7), 425. doi:10.3390/ijgi10070425.
- Di Cecco, G. J., Barve, V., Belitz, M. W., Stucky, B. J., Guralnick, R. P., & Hurlbert, A. H. (2021). Observing the observers: How participants contribute data to iNaturalist and implications for biodiversity science. *BioScience*, 71(11), 1179–1188. <http://bioscience.oxfordjournals.org/>. doi:10.1093/biosci/biab093.
- Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers*, 102(3), 571–590. doi:10.1080/00045608.2011.595657.
- Fast, V. (2024). GIScience I: The rise, fragmentation, and future of VGI. *Progress in Human Geography*. doi:10.1177/03091325241277834.
- Feng, Y. u, Huang, X., & Sester, M. (2022). Extraction and analysis of natural disaster-related VGI from social media: Review, opportunities and challenges. *International Journal of Geographical Information Science*, 36(7), 1275–1316. doi:10.1080/13658816.2022.2048835.
- Fink, D., Auer, T., Johnston, A., Ruiz-Gutierrez, V., Hochachka, W. M., & Kelling, S. (2020). Modeling avian full annual cycle distribution and population trends with citizen science data. *Ecological Applications*, 30(3). [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1939-5582](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1939-5582). doi:10.1002/eap.2056.
- Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M., & Elkinton, J. S. (2009). Observer bias and the detection of low-density populations. *Ecological Applications*, 19(7), 1673–1679. <http://www.esajournals.org/doi/pdf/10.1890/09-0265.1>. doi:10.1890/09-0265.1.

- Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221. doi:10.1007/s10708-007-9111-y.
- Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1, 110–120. doi:10.1016/j.spasta.2012.03.002.
- Haklay, M., Basiouka, S., Antoniou, V., & Ather, A. (2010). How many volunteers does it take to map an area well? the validity of Linus' law to volunteered geographic information. *Cartographic Journal*, 47(4), 315–322. doi:10.1179/000870410 z 12911304958827.
- Haklay, M., & Weber, P. (2008). OpenStreetMap: User-generated street maps. *IEEE Pervasive Computing*, 7(4), 12–18. doi:10.1109/MPRV.2008.80.
- Hu, Y., Goodchild, M., Zhu, A.-X., Yuan, M., Aydin, O., Bhaduri, B., Gao, S., Li, W., Lunga, D., & Newsam, S. (2024). A five-year milestone: Reflections on advances and limitations in GeoAI research. *Annals of GIS*, 30(1), 1–14. doi:10.1080/19475683.2024.2309866.
- Huang, Q., & Wong, D. W. S. (2015). Modeling and visualizing regular human mobility patterns with uncertainty: An example using twitter data. *Annals of the Association of American Geographers*, 105(6), 1179–1197. <http://www.tandfonline.com/toc/raag20/current>. doi:10.1080/00045608.2015.1081120.
- Huang, X., Wang, S., Yang, D. i, Hu, T., Chen, M., Zhang, M., Zhang, G., Biljecki, F., Lu, T., Zou, L., Wu, C. Y. H., Park, Y. M., Li, X., Liu, Y., Fan, H., Mitchell, J., Li, Z., & Hohl, A. (2024). Crowdsourcing geospatial data for earth and human observations: A review. *Journal of Remote Sensing*, 4. doi:10.34133/remotesensing.0105.
- Huang, X., Li, Z., Jiang, Y., Li, X., Porter, D., & Gao, S. (2020). Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLOS ONE*, 15(11), e0241957. doi:10.1371/journal.pone.0241957.
- Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2019). Predicting elections from social media: A three-country, three-method comparative study. *Asian Journal of Communication*, 29(3), 252–273. <http://www.tandf.co.uk/journals/titles/01292986.asp>. doi:10.1080/01292986.2018.1453849.
- Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., Newth, D., & Wu, Y. (2015). Understanding human mobility from Twitter. *PLOS ONE*, 10(7), e0131469. doi:10.1371/journal.pone.0131469.
- Kendall, W. L., Peterjohn, B. G., & Sauer, J. R. (1996). First-time observer effects in the North American breeding bird survey. *Auk*, 113(4), 823–829. <http://www.bioone.org/toc/tauk/132/1>. doi:10.2307/4088860.
- Kottwitz, M., Zhang, G., & Xu, J. (2023). The time- and distance-decay effects of hurricane relevancy on social media: An empirical study of three hurricanes in the United States. *Annals of GIS*, 29(4), 469–484. doi:10.1080/19475683.2023.2236678.
- Kruse, J., Kang, Y., Liu, Y. u-N., Zhang, F., & Gao, S. (2021). Places for play: Understanding human perception of playability in cities using street view images and deep learning. *Computers, Environment and Urban Systems*, 90, 101693. doi:10.1016/j.compenvurbsys.2021.101693.
- Larsen, B. B., Miller, E. C., Rhodes, M. K., & Wiens, J. J. (2017). Inordinate fondness multiplied and redistributed: The number of species on earth and the new pie of life. *Quarterly Review of Biology*, 92(3), 229–265. <http://www.journals.uchicago.edu/doi/pdfplus/10.1086/693564>. doi:10.1086/693564.
- Lee, J., Ho, H., Yoo, S., Kwon, Y., & Sohn, H.-G. (2022). Distant object localization with a single image obtained from a smartphone in an urban environment. *International Journal of Applied Earth Observation and Geoinformation*, 111, 102820. doi:10.1016/j.jag.2022.102820.
- Li, J., Stephens, K. K., Zhu, Y., & Murthy, D. (2019). Using social media to call for help in hurricane harvey: Bonding emotion, culture, and community relationships. *International Journal of Disaster Risk Reduction*, 38, 101212. doi:10.1016/j.ijdr.2019.101212.
- Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of twitter and flickr. *Cartography and Geographic Information Science*, 40(2), 61–77. <http://www.tandfonline.com/doi/pdf/10.1080/15230406.2013.777139>. doi:10.1080/15230406.2013.777139.
- Liu, X., Huang, Q., & Gao, S. (2019). Exploring the uncertainty of activity zone detection using digital footprints with multi-scaled DBSCAN. *International Journal of Geographical Information Science*, 33(6), 1196–1223. doi:10.1080/13658816.2018.1563301.
- Lu, T., Liu, Y., Garcia, A., Wang, M., Li, Y., Bravo-villasenor, G., Campos, K., Xu, J., & Han, B. (2022). Leveraging citizen science and low-cost sensors to characterize air pollution exposure of disadvantaged communities in Southern California. *International Journal of Environmental Research and Public Health*, 19(14), 8777. doi:10.3390/ijerph19148777.
- Malik, M. M., Lamba, H., Nakos, C., & Pfeffer, J. (2015). Population bias in geotagged tweets. In *Proceedings of the international AAAI conference on web and social media*: 9 (pp. 18–27).

- McGough, A., Kavak, H., & Mahabir, R. (2022). In *Revisiting linus' law in OpenStreetMap: An agent-based approach*: 13558 (pp. 123–133). Springer Science and Business Media LLC. doi:10.1007/978-3-031-17114-7_12.
- Mooney, P., & Corcoran, P. (2014). Analysis of interaction and Co-editing patterns amongst OpenStreetMap contributors. *Transactions in GIS*, 18(5), 633–659. doi:10.1111/tgis.12051.
- Neis, P., & 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), 146–165. doi:10.3390/ijgi1020146.
- Sauer, J. R., Peterjohn, B. G., & Link, W. A. (1994). Observer differences in the North American breeding bird survey. *The Auk*, 111(1), 50–62. doi:10.2307/4088504.
- Senaratne, H., Mobasher, A., Ali, A. L., Capineri, C., & Haklay, M. (2017). A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), 139–167. <http://www.tandf.co.uk/journals/titles/13658816.asp>. doi:10.1080/13658816.2016.1189556.
- Solis, P., McCusker, B., Menkiti, N., Cowan, N., & Blevins, C. (2018). Engaging global youth in participatory spatial data creation for the UN sustainable development goals: The case of open mapping for malaria prevention. *Applied Geography*, 98, 143–155. doi:10.1016/j.apgeog.2018.07.013.
- Sui, D., Elwood, S., & Goodchild, M. (2013). *Crowdsourcing geographic knowledge: Volunteered geographic information (VGI) in theory and practice*. Springer. doi:10.1007/978-94-007-4587-2.
- Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10), 2282–2292. doi:10.1016/j.biocon.2009.05.006.
- Unger, S., Rollins, M., Tietz, A., & Dumais, H. (2021). iNaturalist as an engaging tool for identifying organisms in outdoor activities. *Journal of Biological Education*, 55(5), 537–547. doi:10.1080/00219266.2020.1739114.
- Yan, Y., Feng, C.-C., Huang, W., Fan, H., Wang, Y. i-C., & Zipf, A. (2020). Volunteered geographic information research in the first decade: A narrative review of selected journal articles in GIScience. *International Journal of Geographical Information Science*, 34(9), 1765–1791. doi:10.1080/13658816.2020.1730848.
- Zhang, G., Zhu, A. X., Huang, Z. P., & Xiao, W. (2018). A heuristic-based approach to mitigating positional errors in patrol data for species distribution modeling. *Transactions in GIS*, 22(1), 202–216. [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1467-9671](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1467-9671). doi:10.1111/tgis.12303.
- Zhang, G., & Zhu, A. X. (2019a). A representativeness heuristic for mitigating spatial bias in existing soil samples for digital soil mapping. *Geoderma*, 351, 130–143. www.elsevier.com/inca/publications/store/5/0/3/3/3/2. doi:10.1016/j.geoderma.2019.05.024.
- Zhang, G., & Zhu, A. X. (2019b). A representativeness-directed approach to mitigate spatial bias in VGI for the predictive mapping of geographic phenomena. *International Journal of Geographical Information Science*, 33(9), 1873–1893. <http://www.tandf.co.uk/journals/titles/13658816.asp>. doi:10.1080/13658816.2019.1615071.
- Zhang, G., & Zhu, A. X. (2018). The representativeness and spatial bias of volunteered geographic information: A review. *Annals of GIS*, 24(3), 151–162. <http://www.tandfonline.com/toc/tagi20/current>. doi:10.1080/19475683.2018.1501607.
- Zhang, G. (2024). A web-based geovisualization framework for exploratory analysis of individual VGI contributor's participation characteristics. *Cartography and Geographic Information Science*, 1–21. doi:10.1080/15230406.2023.2293891.
- Zhang, G. (2020). Spatial and temporal patterns in volunteer data contribution activities: A case study of eBird. *ISPRS International Journal of Geo-Information*, 9(10), 597. doi:10.3390/ijgi9100597.
- Zhang, G. (2021). Volunteered geographic information. *Geographic Information Science & Technology Body of Knowledge*, 2021(Q1). doi:10.22224/gistbok/2021.1.1.
- Zhang, G., Gong, X. i., & Zhu, D. i (2024a). Geographic proximity and homophily effects drive social interactions within VGI communities: An example of iNaturalist. *International Journal of Digital Earth*, 17(1). doi:10.1080/17538947.2023.2297948.
- Zhang, X., An, J., Zhou, Y., Yang, M., & Zhao, X. (2024b). How sustainable is OpenStreetMap? tracking individual trajectories of editing behavior. *International Journal of Digital Earth*, 17(1). doi:10.1080/17538947.2024.2311320.
- Zhu, A. X., Zhang, G., Wang, W., Xiao, W., Huang, Z. P., Dunzhu, G. S., Ren, G., Qin, C. Z., Yang, L., Pei, T., & Yang, S. (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. <http://www.tandf.co.uk/journals/titles/13658816.asp>. doi:10.1080/13658816.2015.1058387.