Taylor & Francis
Taylor & Francis Group

Cartography and Geographic Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcag20

A web-based geovisualization framework for exploratory analysis of individual VGI contributor's participation characteristics

Guiming Zhang

To cite this article: Guiming Zhang (15 Jan 2024): A web-based geovisualization framework for exploratory analysis of individual VGI contributor's participation characteristics, Cartography and Geographic Information Science, DOI: 10.1080/15230406.2023.2293891

To link to this article: https://doi.org/10.1080/15230406.2023.2293891

Published online: 15 Jan 2024.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tcag20

CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE
https://doi.org/10.1080/15230406.2023.2293891

A web-based geovisualization framework for exploratory analysis of individual VGI contributor's participation characteristics

Guiming Zhang

Department of Geography & the Environment, University of Denver, CO, USA

ABSTRACT

Understanding the participation characteristics of VGI (volunteered geographic information) contributors is important as their active participation and consistent data contribution are key to success of any VGI project. Existing studies on this matter focused primarily on deriving and interpreting participation patterns of contributor groups. There is a lack of investigation into the participation pattern of individual contributors, which can be complementary to a comprehensive understanding of VGI contributors' participation characteristics. Building and using a custom web-based geovisualization framework, this study explores the individual-level participation characteristics of VGI contributors from the perspectives of spatial, temporal, thematic, and social interaction patterns. I conducted geovisual exploratory data analysis on VGI datasets from the iNaturalist biodiversity citizen science project to gain intuitions on the clustering and variabilities of participation patterns in iNaturalist, detect participation pattern shifts over time and form explanation hypotheses, and assess and develop metrics to measure participation. The geovisualization framework is expected to be generally applicable to other VGI datasets for exploring individual-level contributor participation patterns. This work is among the first efforts to explore individual-level VGI participation characteristics through geovisualization and geovisual analytics.

ARTICLE HISTORY
Received 6 January 2023
Accepted 27 November 2023

KEYWORDS
Volunteered geographic information (VGI); volunteer contributor; participation pattern; individual-level; exploratory data analysis; geovisualization and analytics

1. Introduction

The past two decades or so have witnessed booming adoptions of volunteered geographic information (VGI) (Goodchild, 2007) for research and applications in geography, ecology, and sociology, among many other disciplines (Connors et al., 2012; Cui et al., 2021; Yan et al., 2020). Participants in VGI initiatives contribute georeferenced and timestamped observations of the physical and social environments (Zhang, 2021). Such collective voluntary data contribution efforts have produced VGI datasets capable of revealing the dynamics of a wide range of geographic phenomena at various spatiotemporal scales (Fink et al., 2020; Haklay & Weber, 2008; Huang et al., 2020; Zhang, Zhu, Huang, Ren, et al., 2018; Zook et al., 2010). For example, data from the eBird citizen science project were used for modeling avian full annual cycle distribution and population trends at landscape- and regional-scales (Fink et al., 2020).

Volunteer contributors, acting as citizen sensors on the ground, are at the center of VGI. A VGI record typically has four components: "who" (observer) reports "what" (thematic information) at "where" (location) and "when" (time) (Zhang & Zhu, 2018). The "who" component is the basis for other components for it represents the volunteer observer, the subject who conducts observation of a phenomenon of interest at a chosen geographic location and time point. Volunteers are essential to VGI because recruiting and retaining participants to contribute data is key to sustaining any VGI project (Bégin et al., 2018). Moreover, many of the data quality issues associated with VGI can be traced back to the underlying data genesis processes with a focus on the "who" component. Demographic characteristics of VGI contributors may impact VGI data quality. For example, contributors of higher socioeconomic status can afford to travel farther to collect data and elderly contributors are less likely to collect data at places that require strenuous hikes to reach. Spatial, temporal, and thematic biases in VGI are also attributable to the "who" component as it is up to the volunteer to decide what phenomenon to observe and where and when to carry out the observation. As an example, bird watchers who submit data to eBird tend to concentrate birding efforts in populous areas and places of better accessibility and they contribute more birding records during bird migration seasons (Zhang, 2020).

Moreover, social interactions among VGI contributors, whatever form that takes, are intrinsic to VGI

CONTACT Guiming Zhang guiming.zhang@du.edu
© 2023 Cartography and Geographic Information Society

communities (Wang & Ye, 2018). VGI relies on the ability of a social network of contributors to provide enough "eyes" to converge on the facts of the geographic feature under observation, and on a hierarchy of trusted individuals acting as moderators or gate-keepers in the network to assure data quality (Goodchild & Li, 2012). For instance, OpenStreetMap contributors can edit and revise spatial entities digitized by other contributors (Sarkar & Anderson, 2022). iNaturalist citizen science project participants can identify species observations submitted by other participants or vote on identifications proposed by others (Unger et al., 2021). Birding records submitted to eBird, if flagged by the automated filters, are vetted by a network of regional experts in bird occurrences (Kelling et al., 2013). Such social interactions underpinning VGI communities is a key driver of the success of VGI projects (Sbrocchi et al., 2022; Sullivan et al., 2009).

Understanding the participation characteristics of VGI contributors is of significant benefit. To begin with, it sheds light upon designing effective recruiting and engagement strategies targeted at certain types of participants to grow VGI projects and to ensure the longevity of VGI communities (Ponciano & Brasileiro, 2014). As an example, VGI projects can recognize their "loyal" contributors who have been actively and regularly contributing for a long time and facilitate them to mentor newcomers (Aristeidou et al., 2017). Knowledge of VGI contributors' behavior pattern also informs better use of VGI data. Obviously, such knowledge helps increase the awareness of VGI data quality issues that may arise from heterogeneous contributor behaviors (August et al., 2020; Zhang, 2020). One step further, analytical methods that are grounded in a clear understanding of VGI contributors' data contribution characteristics can be devised to effectively account for biases and errors wherever VGI data is used in analysis or modeling exercises (Johnston et al., 2021; Zhang & Zhu, 2019a, 2019b; Zhang, Zhu, Huang, & Xiao, 2018).

Existing studies on participation characteristics in VGI projects and other types of online communities have largely focused on group-level participation patterns, rarely exploring at the individual level. The general methodology adopted by these studies consists of three steps (Aristeidou et al., 2017; Kishimoto & Kobori, 2021; Ponciano & Brasileiro, 2014). First, a set of metrics are computed for individual contributors to depict the temporal (e.g. activity ratio, periodicity), spatial (e.g. active area size), and data content (e.g. proportion of taxa recorded in a biodiversity citizen science project) characteristics of that contributor's activities. Second, based on their scores on the metrics, contributors are clustered into a small number of groups wherein contributors in the same group possess more "similar" values on the metrics than contributors in different groups. Third, the "center" of each group, depicted by the average values on the metrics within that group, is treated as the "typical" activity pattern for that group, which is subsequently interpreted to elucidate group-level participation characteristics. For instance, biodiversity citizen science project participants were clustered into "dabbler", "steady", and "enthusiast" (Boakes et al., 2016), whilst OpenStreetMap contributors were categorized as "nonrecurring mappers", "junior mappers", and "senior mappers" (Neis & Zipf, 2012) or "locals" vs "outsiders" (Quinn, 2016), wherein each of the participant groups represents a distinct participation behavior pattern.

This group-oriented method, albeit instrumental for gaining high-level understanding of the distinct patterns of participation, has certain limitations. First, it reduces the activity patterns of a large number of individual contributors to just a few "typical" group-level profiles. Therefore, anomalous activity patterns (i.e. "outliers") which do not fall well into the groups are simply neglected, although they can be very important for gaining novel insights on abnormal participation behaviors (Xin, 2022). In addition, such profiles represent some sort of "average" participation patterns that cannot be attached to any specific contributor. Thus, interpretations of such "fictitious" profiles may not always be meaningful. Moreover, in some cases, variations among the activity patterns of individual contributors are so subtle that reducing individual's activities to discrete behavioral groups is not always desirable. Instead, it is favorable to simply place individual's activities along axes of the metrics that better reflect the continuous variation between individuals (August et al., 2020). Most importantly, many metrics adopted in the existing studies were derived from conceptual frameworks of user engagement (O'Brien & Toms, 2008), or based on general domain knowledge (August et al., 2020; Boakes et al., 2016). There lacks an emphasis on exploratory data analysis through which effectiveness of the existing metrics can be evaluated or additional metrics developed for depicting participation characteristics in the context of a particular VGI project.

Exploring participation characteristics at the individual level can be complementary to investigating participation patterns at the group level. Individual-level participation profiles, unlike "fictitious" group profiles (i.e. averaged patterns), represent authentic activity pattern of specific VGI participants. Anomalous participants with interesting contribution patterns thus are not excluded from analysis for knowledge discovery (Xin, 2022). Furthermore, profiling the activity pattern of individual contributors is usually an important step in exploratory data analysis, which could be informative to subsequent formal analyses. For instance, by exploring and contrasting individual contributor's spatial, temporal, and thematic (i.e. data content) characteristics, relevant metrics can be developed to measure participation patterns to better reflect the continuous variation in participation behaviors (August et al., 2020), or to differentiate participation patterns for meaningful clustering (Ponciano & Brasileiro, 2014). Lastly, this exploratory data analysis process helps generate hypotheses for explaining the causes of the similarity and difference among participation patterns and any pattern change over time at the individual- or group-levels (Aristeidou et al., 2017; Aristeidou, Herodotou, Ballard, Young, et al., 2021; Kishimoto & Kobori, 2021). In this sense, exploring individual contributor's participation characteristics can be supplementary to group-level analysis in understanding contributor participation patterns. Nonetheless, individual-level participation pattern analysis has been understudied in existing literature.

The examinination of the participation patterns of individual VGI contributors can be well supported by geovisualization and geovisual analytics because VGI data are inherently geospatial (i.e. associated with locations). Exploring individual contributor's participation characteristics, just like any other data exploration processes in general, requires effective tools to visualize the spatial, temporal, and thematic patterns embedded in the data (Nakaya & Yano, 2010). Interactivity between a human analyst and the visualizations is also highly desired as it enables the analyst to flexibly manipulate the visualizations for better sense-making and reasoning (Roth & MacEachren, 2016). Such requirements can be fulfilled by geovisualization and geovisual analytics. Geovisual analytics is defined as "the science of analytical reasoning with spatial information as facilitated by interactive visual interfaces" (Robinson, 2017). Different from traditional geovisualization that emphasizes spatial information visualization, geovisual analytics focuses on analytical reasoning and decision-making while utilizing interactive geovisualization as an important supporting tool (Andrienko et al., 2007; Kraak, 2008).

A variety of geovisualization and geovisual analytics tools have been developed to assist sense-making data in many domains, such as public health (Chen, Roth, et al., 2008), political redistricting (Guo & Jin, 2011), crisis management (Tomaszewski & MacEachren, 2012), vessel movement (Enguehard et al., 2013), and human mobility (Zhang et al., 2019). Specific to VGI-related application of geovisualization and geovisual analytics, it was effectively utilized on social media data to explore public political discourse (Nelson et al., 2015), to analyze crime patterns (White & Roth, 2010), and to examine multi-scale human mobility (Li et al., 2021). A geovisual analytics tool was also designed to help professional users of OpenStreetMap understand contributors' characteristics in specific places (Quinn & MacEachren, 2018). Additionally, geovisual analytics was used to uncover spatio-temporal patterns in event data obtained from RSS (Really Simple Syndication) news feed (Robinson et al., 2017). Geovisualization and geovisual analytics tools generally are highly customized and tailored to the specific datasets and analytical problems at hand since the tools should be designed following "user-centered" principles where much emphasis is placed on assessing key user needs and requirements in the context of a particular application (Robinson, 2017), although a few toolkits were developed to offer generic visualization building blocks based on which customized geovisualization and geovisual analytics tools can be created (Chen, MacEachren, et al., 2008; Ho et al., 2012).

From a technical point of view, early geovisualization tools were mostly implemented as desktop applications (Chen, MacEachren, et al., 2008; Chen, Roth, et al., 2008; Enguehard et al., 2013; Guo & Jin, 2011; Robinson et al., 2017; White & Roth, 2010). In later years, more web-based tools were developed for geovisual analytics (Ho et al., 2012; Li et al., 2021; Nelson et al., 2015; Quinn & MacEachren, 2018; Tomaszewski & MacEachren, 2012; Zhang et al., 2019). The transition toward web-based geovisual analytics, on the one hand, was largely due to the widespread use and easy accessibility of the Internet (Ho et al., 2012). On the other hand, there are increasingly available web mapping tools (e.g. Leaflet, OpenLayers, Kepler.gl), data visualization libraries (e.g. D3.js, Plotly.js), and web development frameworks (e.g. jQuery, Django) that support developing web-based, full-fledged geovisualization and geovisual analytics tools.

For this study, I built and used a custom web-based geovisualization framework (https://guiming.github.io/GeovizVGI/) to explore the participation characteristics of individual VGI contributors along four dimensions of VGI participation characteristics: spatial, temporal, thematic, and social interaction. iNaturalist was used as an example VGI project wherein contributor participation characteristics were explored with the support of the geovisualization framework. Similar to many other citizen science projects, participant contributions in iNaturalist are highly skewed wherein a small number of participants contribute a large portion of the data while a majority of participants contribute very little (Sauermanna & Franzonib, 2015). The contributions can also be very uneven across space, time, and species taxonomic categories, resulting in spatial, temporal, and thematic biases (DiCecco et al., 2021; Zhang, 2020, 2022). Previous studies attempted to examine iNaturalist contributors' participation pattern at the group level. For instance, Tupikina et al. (2021) categorized iNaturalist contributors as "low activity", "observers", "identifiers", or "high activity" based on the number of observations and identifications they contribute. Kishimoto and Kobori (2021) grouped iNaturalist contributors into "enthusiastic", "off-and-on", "temporary", or "intense" according to contributor's activity ratio, daily observation, and daily devoted time. The distinct participation characteristics of young iNaturalist contributors against other age groups were also analyzed (Aristeidou, Herodotou, Ballard, Higgins, et al., 2021; Aristeidou, Herodotou, Ballard, Young, et al., 2021). However, all these works stop short of exploring iNaturalist participant behavior at the individual level. Furthermore, geographic components of contributor behavior were not a focus of these studies, and only a few metrics were used in the studies while all four dimensions of participation characteristics are worth considering. Particularly, existing studies exploring VGI participation characteristics mostly ignored the social interaction dimension, probably due to the lack of data capturing the interactions. Nonetheless, understanding the social interaction dimension is as important as the other three dimensions because social interactions are intrinsic to many VGI communities.

Compared to visualizations available on existing VGI platforms that are often limited to certain aspects of contributions (e.g. a map of observations on iNaturalist website), the proposed geovisualization framework provides much richer visualizations (e.g. interactive maps, charts, and graphics) for exploring individual contributor's participation patterns along the spatial, temporal, thematic, and social interaction dimensions. To the best of my knowledge, this study is among the first efforts to explore the participation characteristics of individual VGI contributors using custom geovisualization and geovisual analytics tools (alongside Quinn & MacEachren, 2018). Section 2 presents the datasets used in this study. Section 3 details the design and implementation of the geovisualization framework. Section 4 demonstrates the utility of the framework through application scenarios related to exploring VGI contributor's participation characteristics, followed by discussion and conclusion in Section 5 and Section 6, respectively.

2. Datasets

iNaturalist, launched in 2008, is among the world's largest geospatial biodiversity citizen science projects. It is a social network and community of biologists, naturalists, and citizen scientists who contribute, share, and identify species sightings across the globe through the iNaturalist website or mobile app (Unger et al., 2021). As of December 2022, iNaturalist has compiled over 122 million observations on more than 404,500 species based on contributions from nearly 2.5 million observers and over 281,400 identifiers (iNaturalist, 2022b). This study uses iNaturalist data from 2019 to 2020 to explore iNaturalist contributors' participation patterns in these two years.

Each iNaturalist observation contains information regarding "who" observed "what" species at "where" and "when" and, if applicable, identified by "whom." Besides the spatial, temporal, and thematic information embedded in iNaturalist data, the social interactions therein (i.e. species identification interactions) are also of interest when examining the participation characteristics of iNaturalist contributors. iNaturalist data were obtained from two sources: raw observations downloaded from iNaturalist ($n = 43{,}033{,}502$ raw observations were contributed by 1,454,501 observers in 2019 and 2020) (iNaturalist, 2022c), and research-grade observations provided by the Global Biodiversity Information Facility (GBIF) containing only observations that meet established quality control requirements and were published under certain licenses or waivers ($n = 17{,}909{,}484$ research-grade observations from 435,802 observers were identified by 326,500 identifiers in 2019 and 2020) (Ueda, 2022). Each raw observation includes geographic coordinates of the location of observation, observation date and time, species taxon (if any), and observer information (e.g. user id, login). A research-grade observation in addition has information on the identifier and time of identification, which is crucial for reconstructing a social network representing species identification interactions among iNaturalist contributors.

Additionally, the land cover type at each iNaturalist observation location was extracted from the yearly MODIS global land cover type dataset (500 m resolution; 2001–2019) (Sulla-Menashe & Friedl, 2018). iNaturalist observation locations in 2019 and 2020 were assigned land cover type based on the most recent 2019 land cover data. Integrating land cover data to provide contextual information about the environment being observed by VGI contributors is a novel piece of this work (i.e. few prior studies analyzing VGI data had integrated land cover data), which can increase the potential utility of the custom geovisualization framework. Exploring contribution patterns in iNaturalist requires geovisualization and geovisual analytics as iNaturalist data are all linked to places (i.e. observation locations) and landcover at those places is integrated into the system.

3. The geovisualization framework

3.1. Usage scenarios analysis

Target users of the geovisualization framework developed in this study (https://guiming.github.io/GeovizVGI/) are researchers who, acting as human analysts, use the geovisualization tool to conduct exploratory data analysis (Kraak, 2003) on VGI datasets to examine the data contribution behavioral pattern of individual VGI contributors (Quinn & MacEachren, 2018). Such explorations are to support application scenarios involving three inter-connected tasks: 1) discovering the clustering and variabilities in VGI contributor's participation patterns (Ponciano & Brasileiro, 2014), 2) detecting VGI contribution pattern change over time and generating explanation hypotheses to explain pattern change (Kishimoto & Kobori, 2021), and 3) assessing the utility of existing metrics and developing additional ones to characterize and differentiate participation patterns (August et al., 2020). Information needed for these tasks is all embedded in VGI datasets as a VGI record often contains spatial, temporal, thematic (data content), and social interaction information (Zhang & Zhu, 2018).

Nonetheless, it can be challenging to achieve the tasks by ad-hoc exploration of individual contributor's participation patterns. To enable more guided exploration, metrics (features) reflecting contributor participation characteristics were derived for clustering contributor participation patterns into groups (i.e. number of species observations, number of active months, standard distance of observation locations, number of species taxonomic kingdoms, number of land cover types across observation locations, and number of other contributors whom a contributor had interactions with through species identification). The clustering was done off-line during data pre-processing using the k-mean clustering algorithm wherein the number of clusters ($n = 4$) was determined based on elbow method (Kodinariya & Makwana, 2013) (see Supplemental Material S2 for more details). Clustering contributor participation patterns beforehand facilitates the analytics tasks by allowing 1) the examination of the clustering of participation patterns, 2) the comparison of individual participation patterns against cluster centers (i.e. "average" patterns) to reveal participation variabilities (e.g. typical vs. atypical contributors in each cluster), and 3) querying by directional changes in the patterns of contributor's participation (e.g. examining contributor whose participation had increased from 2019 to 2020).

The geovisualization tool shall support the human analyst's exploration of participation patterns not only at the individual level but also at the complementary group level. Specifically, it needs to be able to retrieve information regarding the participation pattern clusters and VGI data entries contributed by a specified contributor over a specified timeframe. It should then construct visualizations to present the spatial, temporal, and thematic distributions of the data, and the social interactions involving the focal contributor. The analyst is offered ample opportunities to interact with and manipulate the visualizations to better suit the analyst's cognitive processes. The analyst forms intuitions, based on the sense-making of the visualizations, to answer questions related to the application scenarios. The questions may be regarding both individual- and group-level participation characteristics and the involvement of multiple dimensions of contribution characteristics (Table 1). Answers to these and similar questions explicate the spatial, temporal, thematic and social interaction patterns depicting the participation characteristics of the contributor. The analyst repeats this process to explore the participation characteristics of many different contributors or of the same contributor in different time periods with reference to participation cluster centers. Intuitions and insights gained in the geovisualization-aided exploratory data analyses would facilitate the aforementioned analytics tasks.

3.2. System design

For easy accessibility to potential users, the geovisualization framework was designed following a web-based client-server architecture (Figure 1). Users (human analysts) rely on a web browser (client) to access the geovisualization tool (i.e. a web page) published on a web server. The web server is connected to application servers that host applications handling queries to the underlying database server where the VGI dataset resides. The web server can also request base map web map services (WMS) from third-party servers as a backdrop for geovisualization.

When an analyst lands on the web page provisioning the geovisualization tool, it presents a default base map (with basic map interaction options such as pan, zoom, layer switch, etc.) delivered from third-party web map services through the web server. A visual interface is also provided for the analyst to set database query parameters (e.g. contributor identifier, time frame). The query parameters and additional parameters derived from the base map display (e.g. spatial extent, zoom level) are sent to the web server as requests. The web server then processes and forwards requests containing the parameters to the application servers. The application servers hosting multiple geospatial feature services and tile services subsequently map the requests into database queries written in SQL (structured query

Table 1. Example questions posed under the three application scenarios regarding participation characteristics at different levels (group vs. individual) and along various dimensions (spatial, temporal, thematic, and social).

Scenario #1: Discovering the clustering and variabilities in VGI contributor's participation patterns.		
Does a contributor conduct observations in a small (or large) geographic area?	Individual	Spatial
Does a contributor submit data regularly or just sporadically? Do the submissions tend to be geographically clustered or spread out evenly?	Individual	Temporal and spatial
Does a contributor submit many different (or just a few) themes of data? Are submissions from many different places (e.g. countries, states, counties)?	Individual	Thematic and spatial
Does a contributor interact a lot (or very little) with others in the community? Does a contributor interact more with those who are geographically closer?	Individual	Social and spatial
How many distinct participation clusters exist among the contributors?	Group	All four
What are the typical characteristics of each participation cluster?	Group	All four
Are there anomalous contributors in each cluster? What are the participation characteristics of such an anomalous contributor?	Individual	All four
Scenario #2: Detecting VGI contribution pattern change over time and generating explanation hypotheses.		
Does a contributor conduct observations in a smaller (or larger) geographic area?	Individual	Spatial
Does a contributor submit data more (or less) frequently? Do those who contribute more (or less) frequently tend to concentrate in certain geographic region?	Individual and group	Temporal and spatial
Does a contributor's interested data themes grow (or shrink)?	Individual	Thematic
Does a contributor's interaction with others increase (or decrease)? Is the increase (or decrease) more (or less) significant between others who are geographically closer (or farther apart)?	Individual	Social and spatial
Who are the contributors' whose participation has increased (or decreased)?	Group	All four
What are the main characteristics of such a pattern change?	Group	All four
What hypotheses could be generated to explain such a pattern change?	Group	All four
Scenario #3: Assessing metrics to measure VGI participation characteristics and developing new metrics.		
Does *standard distance* of observations differentiate the mobility of two contributors? What new spatial metrics may be used to measure participation (e.g. number of countries across observations)?	Individual	Spatial
Can *daily devoted time* be adopted to measure participation? What new temporal metrics may be used to measure participation (e.g. median weekly observations)?	Individual	Temporal
Does the number of species indicate the breath of a contributor's interests? What new thematic metrics may be used to measure participation (e.g. species taxonomic frequency distribution at the kingdom level)?	Individual	Thematic
Can node degree (number of contributors interacted with) reflect the social interaction of a contributor? What new social metrics may be used to measure participation (e.g. PageRank score)?	Individual	Social

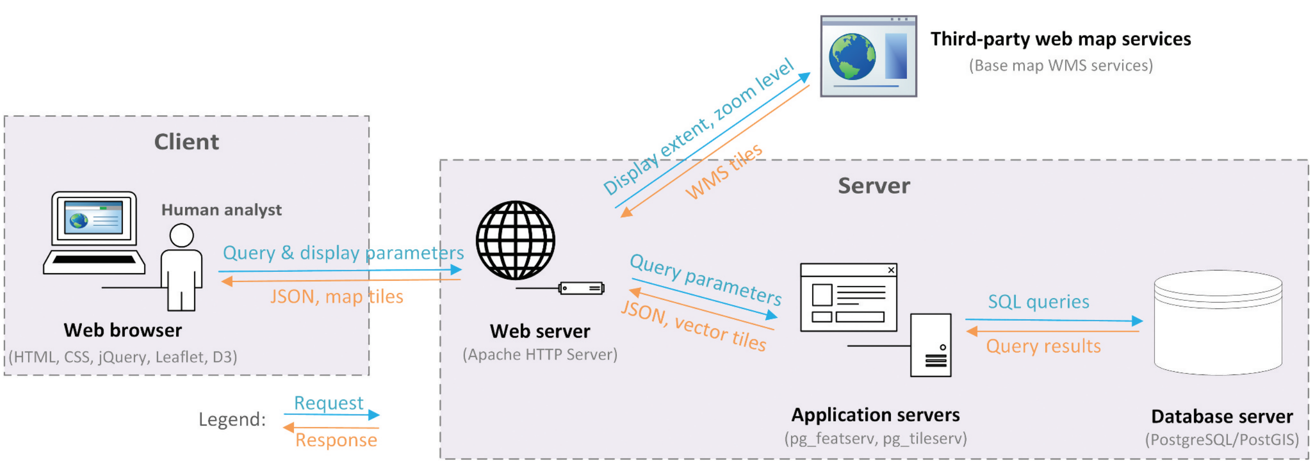

Figure 1. The client-server system architecture of the web-based geovisualization framework.

language), which are eventually executed on the database server to retrieve contributor-specific VGI data entries and to compute auxiliary statistics.

Query results (data entries and statistics) are returned from the database server to the application servers where they are encoded as JSON (JavaScript object notation) data objects or rendered as vector tiles to be delivered back to the web server. The JSON objects and vector tiles are sent to the browser as responses. Finally, in the browser, map tiles are displayed on top of the base map and various visualizations are constructed based on data contained in the JSON objects. The user interacts with the map and the visualizations to explore participation characteristics of a focal individual contributor with reference to participation cluster centers. The user can update query parameters to retrieve data for a different contributor to compare their contribution patterns, or on the same contributor but in different time frames to compare their patterns at different times. The user also has the option to query contributors by participation pattern change direction (e.g. contributors whose contribution pattern were in cluster 2 in 2019 but moved up to cluster 3 in 2020). The map and visualizations are updated automatically upon receiving query result from the server. See Supplemental Material S1 for detailed technical implementation of the geovisualization tool.

3.3. Geovisualization interfaces

The basic layout of the geovisualization tool is shown in Figure 2 (with visualizations for an example focal contributor). The base map with standard interactivity functionalities (zoom, pan, layer switch, layer visibility control, etc.) provides a background for visualizing participation patterns. A panel floating atop the base map is presented to the human analyst for customizing database query parameters, including from and to dates that define the timespan, number of contributors to randomly select from all contributors (stratified by participation cluster), number of typical and atypical contributors to randomly select from each participation cluster and, optionally, a list of user-specified contributors. Here typical and atypical contributors are defined as contributors whose feature distance to their respective cluster center (as measured by the six features used for clustering) is within the lower 5 percentile and beyond the 95 percentile, respectively. Examining typical and atypical contributors offers the analyst opportunities to explore both "average" (typical) and anomalous (atypical) participation patterns. The analyst can also choose to query by contribution pattern change direction over the years. These parameters are used to query and load the six metrics (features) of cluster centers and an initial batch of contributors. Metrics of cluster centers and individual contributors are then visualized on a parallel coordinate plot so a particular contributor's participation pattern can be compared to the respective cluster center. The user can modify any of the query parameters to reload another batch of contributions or choose a particular contributor (the focal contributor) from the loaded batch for contributor-specific queries and in-depth geovisual explorations.

The returned species observations contributed by the focal contributor are mapped on top of the base map. A control panel provides access to additional visual interfaces constructed based on other data retrieved from the server (e.g. time series, identification interactions) to facilitate exploring participation characteristics of the focal contributor along the spatial, temporal, thematic, and social dimensions, as described in detail below.

3.3.1. Spatial characteristics

Spatial patterns of the focal contributor's observations are visualized in multiple ways. A point layer rendering the observation locations is overlaid on the base map to show the spatial distribution of observations. Each point is clickable to activate a popup window displaying detailed information regarding an observation (Figure 2). A heatmap is created from the observation locations to highlight spatial clustering of the observations. The "spatial" checkbox on the control panel displays additional interactive pie charts to present distribution of the observations over land cover types, countries, and sub-country administrative units (e.g. states, provinces) (Figure 3). These interactive visual constructs are intended to assist the analyst's exploration of the spatial characteristics of VGI contributor's participation.

3.3.2. Temporal characteristics

Temporal trends in the focal contributor's contributions are presented as bar-chart time series. The analyst can switch between three target variables to be visualized on the time series, namely number of observations, number of identifications (among research-grade observations), and standard distance of the observation locations (an indicator of the spatial spread of the contributor's observation activities) (Figure 4). For each time series, the analyst can switch between day, month, and year to display temporal variability in the target variable at different time granularities. Time series of each variable includes two views, a zoomable overview bar chart at the bottom, and a zoom-in bar chart on the top. The analyst can interactively select a focus time period by "brushing" or panning the time axis of the overview, and the zoom-in view automatically updates to show temporal variations in the target variable during the selected period. The analyst can also choose to update the background map on-the-fly to only show the focal contributor's observations within the selected period. These interactive visualizations allow the analyst to extract temporal patterns of VGI contributor's participation.

3.3.3. Thematic characteristics

Thematic patterns in the focal contributor's contributions are concerned primarily with taxonomic distribution of the observations (Figure 5). An interactive pie chart shows frequency distribution of species observations at the kingdom taxon level. Down to the species level, word clouds are utilized to depict the frequency distribution of the observed species common names and scientific names. Through the pie chart and word clouds, the analyst gains intuitions on the thematic characteristics of the data contributed by the VGI contributor.

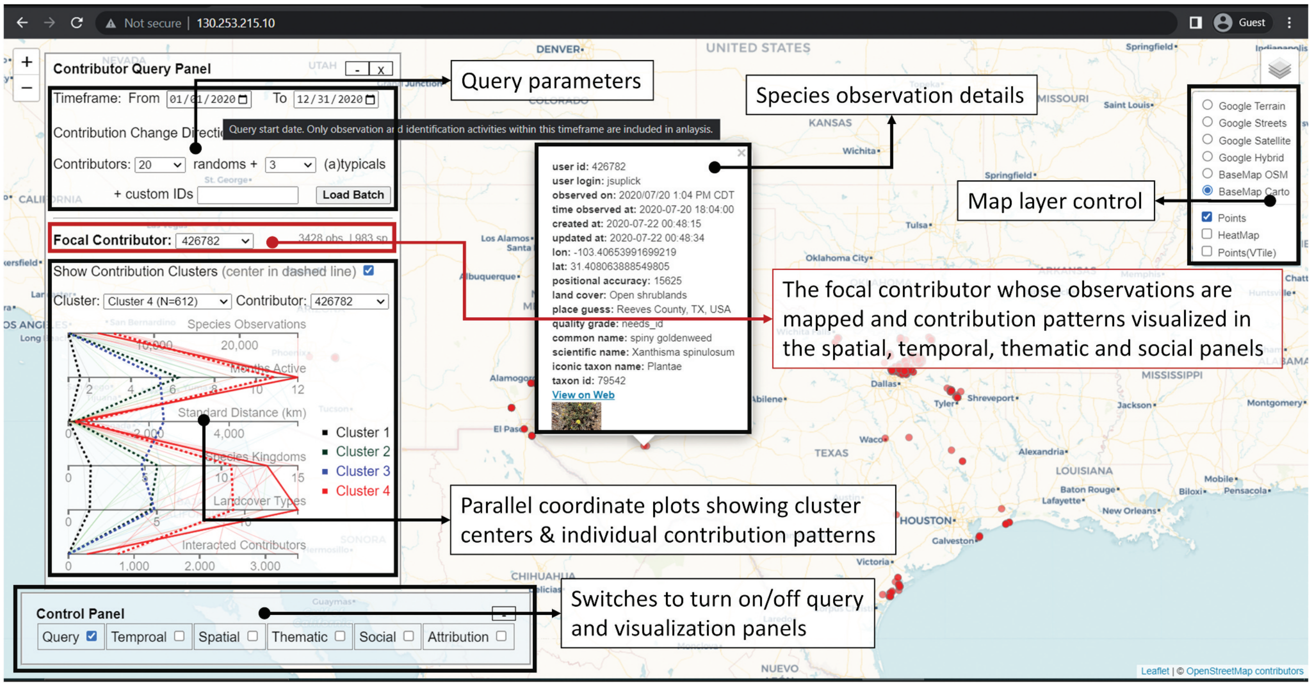

Figure 2. Basic layout of the geovisualization interface for exploring individual VGI contributor's participation characteristics. A mouse tooltip explanation is provided for each query and visualization parameter on the query panel (textbox, drop-down, checkbox, button, etc.) to explain what it is and what it controls (e.g. explanation for the start date is shown). On the parallel coordinate plot, cluster centers are in dashed lines, individual contributors in solid lines, and the focal contributor is in the highlighted solid line. Lines are drawn based on values of the six features reflecting contribution characteristics: number of species observations, number of active months, standard distance of observation locations, number of species taxonomic kingdoms, number of land cover types across observation locations, and number of other contributors whom a contributor had interactions with through species identification.

3.3.4. Social characteristics

Social interactions between the focal contributor and other VGI contributors are captured with an identification social network wherein nodes are contributors who either identified the focal contributor's observations or had their observations identified by the focal contributor (Figure 6). The node at the center of the network represent the focal contributor. An edge represents species identification interactions between two contributors. Nodes are color-coded by contributor id and node size is proportional to the number of identifications made by the respective contributor. Edge width is determined based on the number of identifications between two nodes (link weight). Options are offered to interactively manipulate (e.g. dragging node or edge, viewing node or edge information, opening user profile page on iNaturalist site, etc.) or simplify the social network (e.g. by increasing link weight threshold). The interactive social network visualization enables the analyst to explore patterns in the social interactions involving the focal contributor. These simple network visualization and analysis options provide a starting point for exploring VGI contributor's social interaction characteristics, although more advanced visual network analytics (Arleo et al., 2022) could be incorporated in future iterations of the geovisualization tool.

4. Demonstration of application scenarios

To demonstrate the utility of the geovisualization tool, it was utilized by the author (referred to as the analyst) to explore participation patterns of individual contributors with reference to participation cluster centers by examining the various (geo)visualizations that present the spatial, temporal, thematic distributions of the data contributed by the contributor and the social interactions involving the contributor. Such explorations are to support the three intended tasks (see Section 3.1): 1) examining the clustering and variabilities of iNaturalist contributors' participation patterns, 2) detecting contribution pattern change over time and forming hypotheses to explain such pattern change, and 3) assessing the utility of existing metrics and developing additional ones to characterize and differentiate participation patterns.

4.1. Clustering and variabilities of participation characteristics

4.1.1. Participation pattern clusters

The clustering of all iNaturalist contributors' participation patterns in 2019 and 2020 (Section 3.1) resulted in four clusters with increasing level of activeness in participation

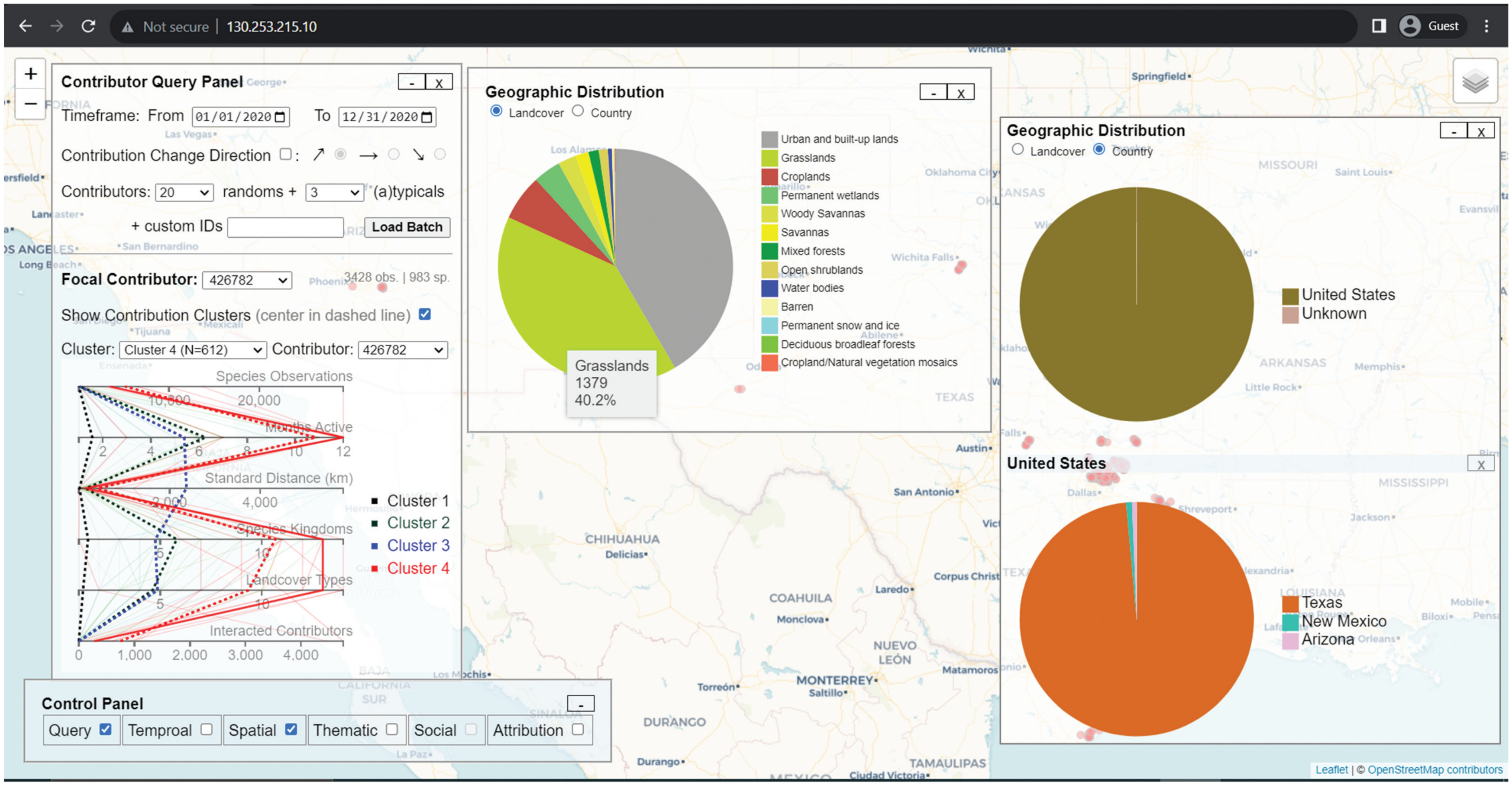

Figure 3. Visualizations of the distribution of VGI observations across geographic space (background point location map) and across land cover types, and countries and sub-country administrative units (pie charts).

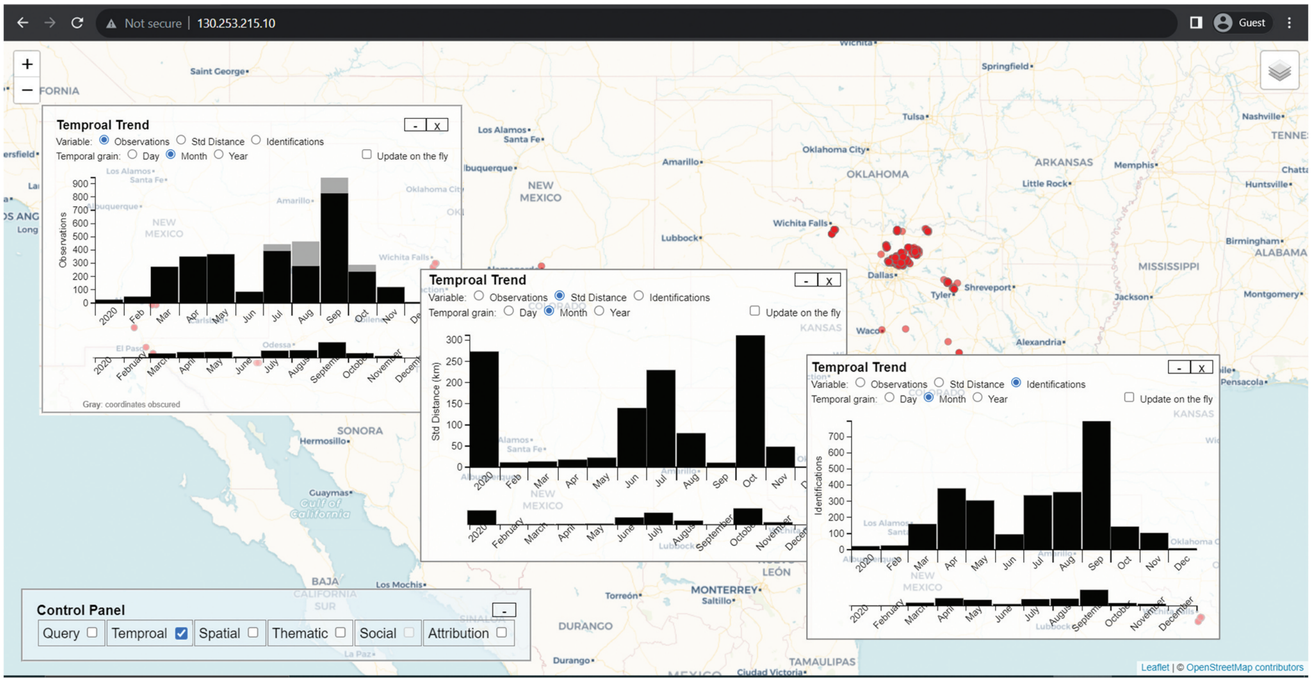

Figure 4. Visualizations of temporal pattern in three target variables (number of observations, standard distance of the observation locations, and number of identifications) through bar-chart time series.

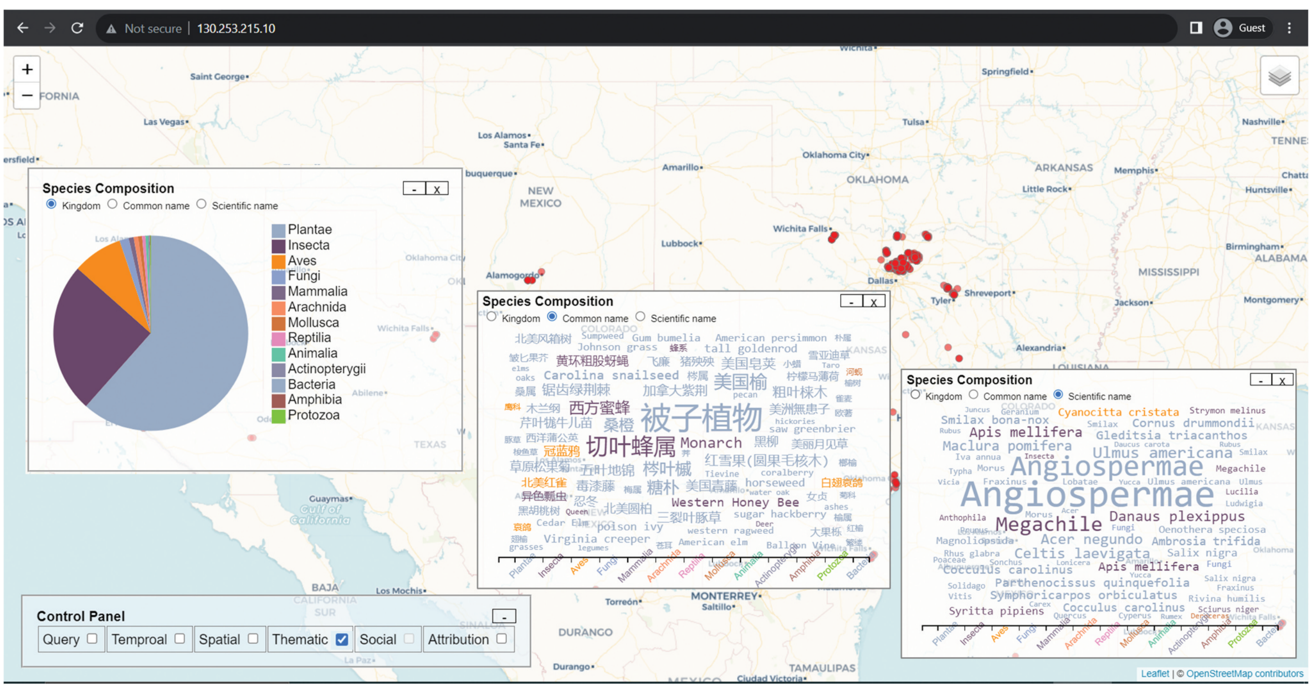

Figure 5. Visualizations of patterns in data content of the focal contributor's observations in terms of the frequency distribution of species observations at the kingdom taxon level (pie chart) and at the species level (word clouds based on frequencies of species common name and scientific name).

(Table 2). The quality of the clustering result is satisfactory as reflected by an average Silhouette score of 0.65, indicating that a reasonable clustering structure was found (Ponciano & Brasileiro, 2014) (see Supplemental Materials S2 for details). Cluster 1 represents the majority of contributors (~84%) who were least active in contributing and, notably, they often did not engage with other contributors through species identification interactions.

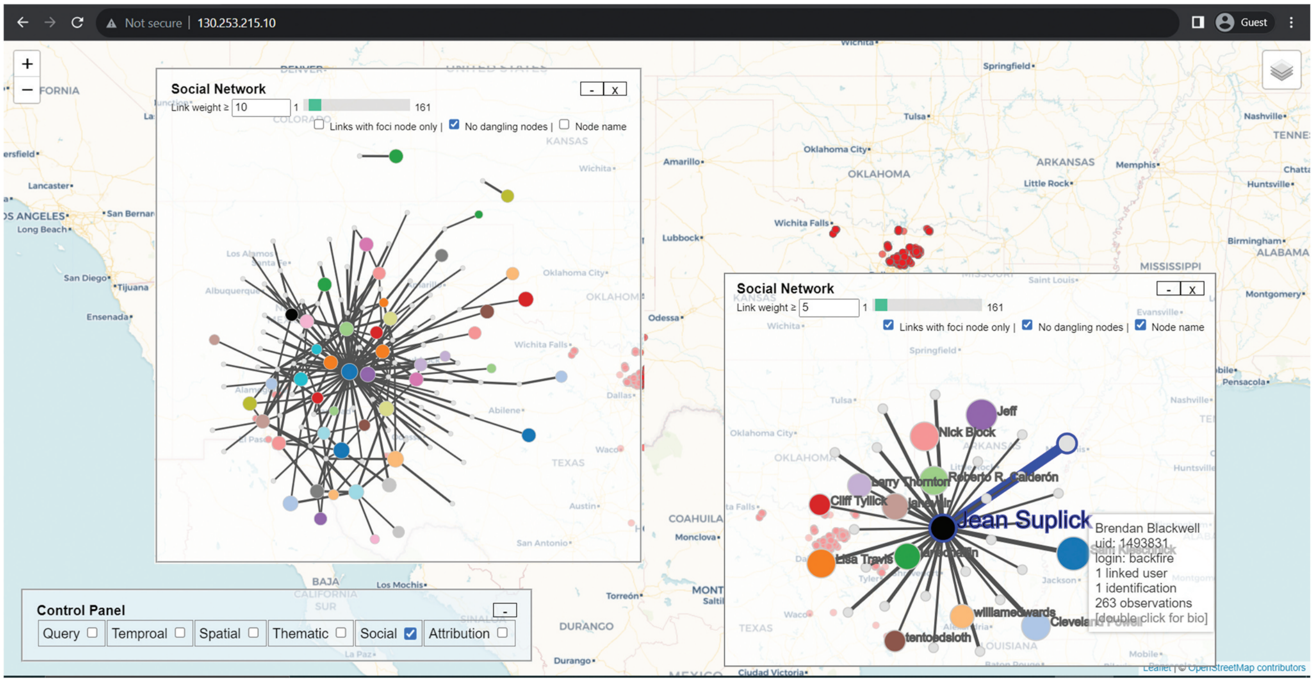

Figure 6. Visualizations of social interaction patterns of the focal contributor through a social network constructed based on species identification interactions between the focal contributor (at the center of the network) and other contributors.

Contributors in cluster 2 (~15%) and cluster 3 (~1%) both had moderate-level contributions and limited interaction with others. However, contributors in cluster 3 had a much larger standard distance, indicating a much higher mobility level (e.g. traveling abroad to make observations). Contributors in cluster 4 (~0.05%) were highly active, submitting thousands of species observations and interacting with hundreds of contributors.

The above contribution cluster centers reflect only some sort of abstract "average" pattern in each contributor group. They do not fully represent the variability of participation characteristics, even within the same cluster. The geovisualization tool offers the analyst a variety of interactive visualizations to explore individual participation characteristics by revealing the spatial, temporal, thematic, and social patterns embedded in the data contributed by individual VGI contributors (August et al., 2020). The analyst therefore is able to not only concretize the typical contribution patterns (as measured by the six features used for clustering) for the clusters, but also investigate contributors with non-typical (anomalous) contribution patterns to explore variabilities therein. Presented below are the participation profiles of typical contributors from each cluster (i.e. randomly selected contributors whose feature distance to the respective cluster center was within the lower 5 percentile) (Figure 7). These typical participation profiles were used here as examples to visually contrast the distinct participation characteristics that exist among participants in the iNaturalist community.

The contributor typical of cluster 4 (highly active contributors) regularly contributes large amounts of data. This contributor was active every month of 2019 and submitted over 3,200 observations (the standard distance among the observation locations was 797.9 km) from habitats of 12 different landcover types on more than 1,200 species in 13 taxonomic kingdoms. The contributor was socially active, having interacted with 1,035 other iNaturalist contributors and identified 2,129 species observations. The geovisualization framework revealed compositions of this particular

Table 2. Cluster centers of iNaturalist participant contribution patterns.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
# Species observations	5.3	108.7	95.5	5149.0
# Active months	1.6	6.3	5.4	10.9
Standard distance (km)	16.6	154.5	2373.5	414.7
# Species kingdoms	1.5	5.8	4.8	10.7
# Land cover types	1.2	4.6	4.9	9.3
# Interacting contributors	0.0	2.1	2.1	749.9
# Contributors in 2019	586,091	96,304	15,297	318
# Contributors in 2020	871,196	164,655	7,048	612

contributor's contributions (which may not be representative of other contributors in cluster 4): The observations were made from four countries (90.1% observations in US); Most species observations were in deciduous broadleaf forests (46.1%), followed by woody savannas (19.8%), and urban and built-up lands (17.4%); More contributions were made in the spring and summer months; Most observations were on plants (51.3%), followed by insects (30.7%).

The contributor representing cluster 3 (moderately active contributors with high mobility) had much fewer contributions and was active during only five months in 2019. Throughout 2019, this contributor submitted only 41 observations (the standard distance was 2332.5 km) from seven landcover types on 27 species in five taxonomic kingdoms and identified 9 species observations from seven other contributors. The contributor typical of cluster 2 (moderately active contributors) submitted 50 observations (the standard distance was 25.4 km) on 49 species in 7 kingdoms and identified 34 observations from three other contributors (Figure 7). Overall, this contributor had a contribution level similar to the cluster 3 contributor above, with a major difference that species observations are much more localized (a much smaller standard distance compared to cluster 3). Finally, the contributor typical of cluster 1 (least active contributors) made only sporadic data contributions (Figure 7). This contributor contributed only four species observations in just three days and did not identify any species records.

4.1.2. Participation pattern variabilities

The typical participation profiles of the four clusters do not represent the full breath of participation variabilities among the contributors, even within the same participation cluster. The analyst can use the geovisualization tool to explore participation variabilities across contributors by loading and examining a sample of both typical and atypical contributors in each cluster (contributors whose feature distance to their respective cluster center is within the lower

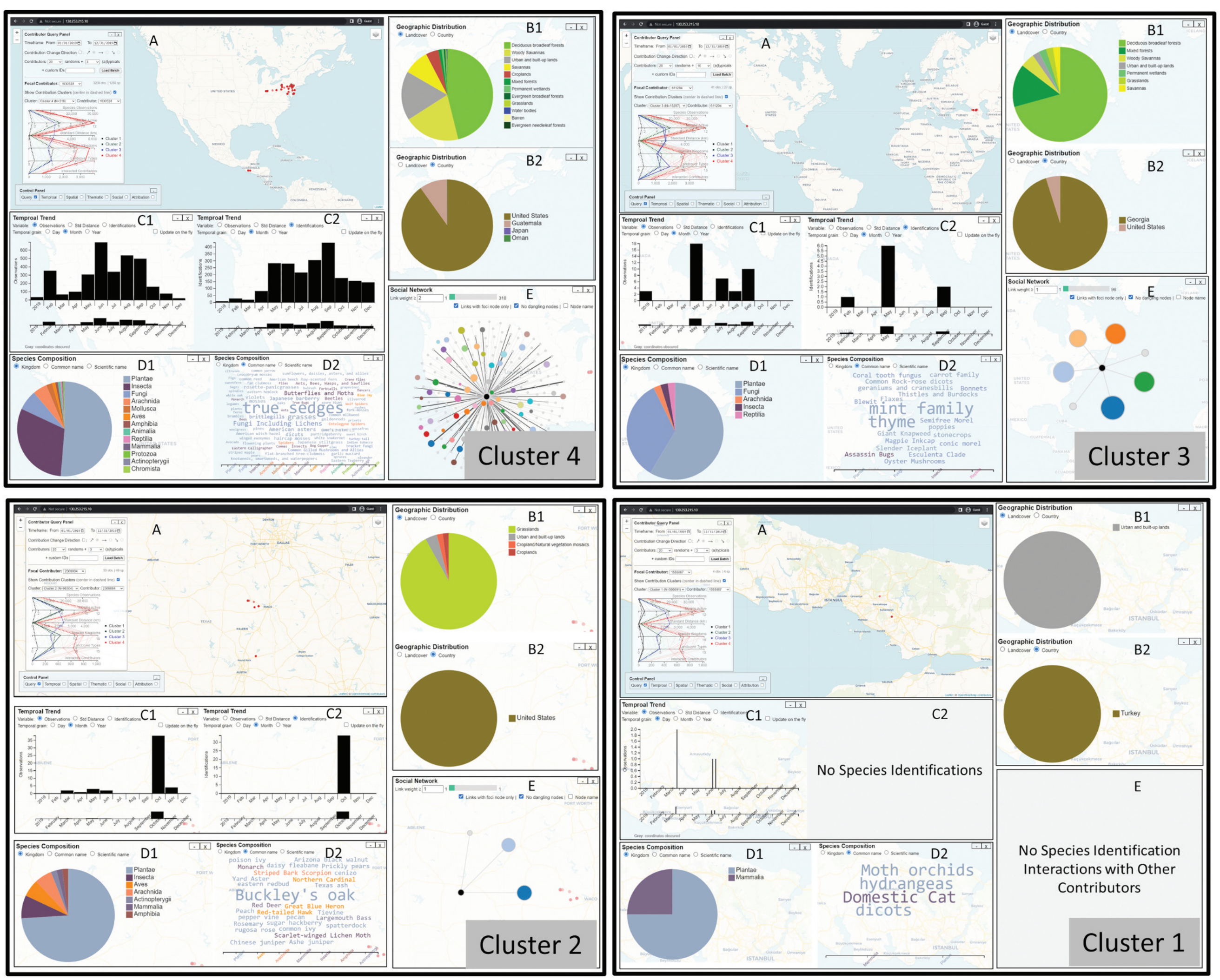

Figure 7. Participation profiles of contributors typical of the four clusters.

5 percentile and beyond the 95 percentile, respectively; see Section 3.3). For instance, a contributor in cluster 4 (highly active) in 2019 submitted over 37,000 species observations on more than 5,200 species in 13 taxonomic kingdoms and on 15 land cover types across 59 countries (Figure 8). This anomalous contributor had a much higher level of participation compared to a typical contributor in that cluster in terms of the usually large amount of species observations and extensive breadth of taxonomic and geographic coverages. Another abnormal contributor in this group (Figure 8) submitted average amount of species observations but had identified observations contributed by over 4000 other contributors in the community in 2019, an astonishingly higher level of social interaction compared to a typical contributor in cluster 4.

4.2. Change in participation patterns

4.2.1. Contribution pattern change

In this study, based on the clustering results, a contributor's participation pattern was regarded as "changed" if the contributor belonged to different participation clusters in 2019 and 2020. Table 3 shows the trend in contribution pattern change over the two years. Among the 287,021 iNaturalist contributors who contributed data in both years 43,886 (15.3%) had increased participation 36,468 (12.7%) had decreased participation, and the remaining (72%) stayed at the same participation level.

Using the geovisualization tool, the analyst can query contributors by participation pattern change direction, who can then compare participation characteristics of the same contributor in different periods to verify and understand participation pattern change over time (Bégin et al., 2018; Kishimoto & Kobori, 2021). As an example, an iNaturalist contributor obtained through query by decreasing contribution pattern (from cluster 4 in 2019 to cluster 2 in 2020) was identified and analyzed in detail. Although this particular example of pattern change by no means exhausts all participation pattern changes in the iNaturalist community, it

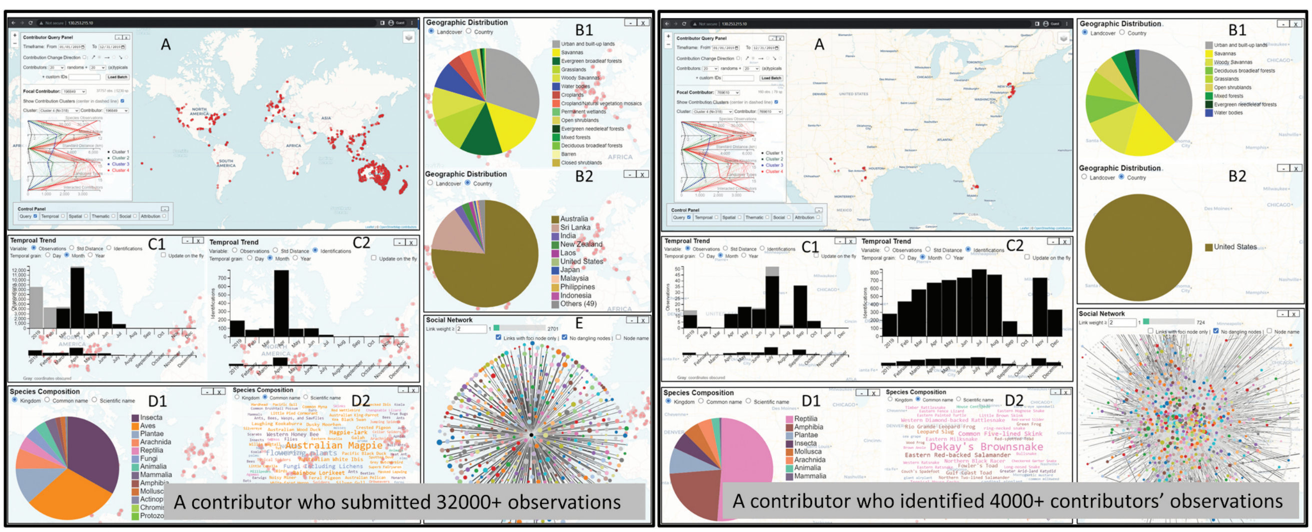

Figure 8. Participation profiles of two anomalous (atypical) contributors.

Table 3. The number of iNaturalist contributors in each category of participation pattern changes from 2019 to 2020. Cells on the diagonal (in gray) represent contributors whose participation stayed at the same level. Above- and below-diagonal cells (in light green and light orange) represent contributors whose participation increased and decreased, respectively.

2019 \ 2020	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	152,279	40,281	1,774	21
Cluster 2	25,625	52,723	1,476	284
Cluster 3	4,560	6,217	1,415	50
Cluster 4	6	57	3	250

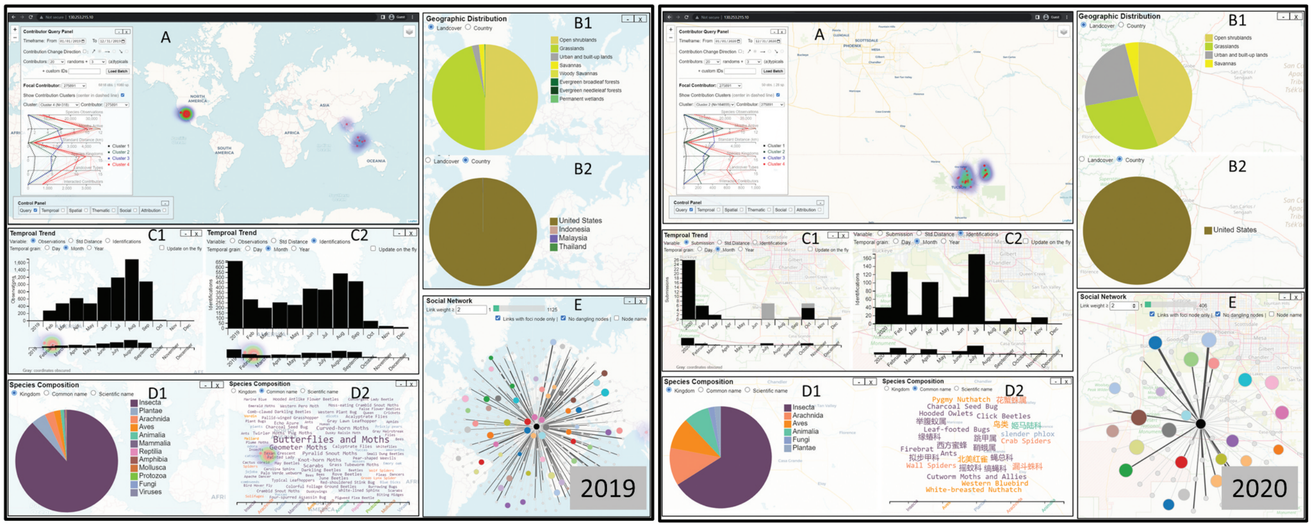

Figure 9. Participation patterns of the same contributor in 2019 and 2020.

exemplifies how the geovisualization tool could be adopted to examine participation pattern change.

Overall, this contributor submitted much less data in 2020 (i.e. 50 observations on 45 species) compared to 2019 (i.e. 6,818 observations on 2,598 species) (Figure 9). Geographically, observations in 2020 were all from a relatively small area in Tucson, US whilst observations in 2019 were from four different countries (the majority of observations were in Tucson, US). In terms of land cover, the percentage of observations on urban and built-up lands increased from 2.2% (2019) to 24% (2020). Temporally, monthly contributions in 2019 were more consistent while contributions in 2020 were of much smaller numbers and were more variable. There were also more observations reported with obscured geographic coordinates in 2020, indicating increased geoprivacy awareness (Figure 9.C1). With respect to observed species composition, the share of arachnids (e.g. spiders, scorpions) increased from 2.2% (2019) to 10% (2020), although insects and animals remained to be the prominent taxa in both years. In terms of community social interactions, the contributor interacted with much fewer contributors and identified fewer observations in 2020 (i.e. 350 connected contributors; 583 identifications) compared to 2019 (i.e. 923 connected contributors; 3495 identifications). The above pattern changes all indicate reduced participation of this particular contributor in 2020.

4.2.2. Hypotheses to explain pattern change

The analyst can examine multiple contributors sampled from any particular pattern change category (Table 3) to visually investigate how a contributor's participation pattern changed along the spatial, temporal, thematic, and social dimensions. The analyst thus can gain intuitions on the commonalities across contributors in that pattern change category, which would inform generating plausible hypotheses to explain the observed pattern change. For example, participation pattern changes along the four dimensions in Figure 9 all indicate reduced participation of this contributor in 2020. The analyst investigated other contributors' participation patterns whose participation had similarly decreased (from cluster 4 in 2019 to cluster 2 in 2020) and confirmed that this change trend was also present among other iNaturalist contributors. Among many potential explanations, the analyst came up with one reasonable hypothesis explaining this pattern change, which is as follows. This contributor used to carry out many observations in natural environments (only 2.2% of the observations in 2019 were on urban and built-up lands), sometimes even at sites in foreign countries. However, due to mobility restrictions and health risks brought by the COVID-19 (coronavirus disease 2019) pandemic (Kishimoto & Kobori, 2021), this participant had to stay in the hometown throughout much of 2020 and contributed a smaller number of observations covering fewer taxonomic kingdoms and fewer land cover types. Meanwhile, the percentage of observations made on urban and built-up lands had increased to 24%, a trend consistent with documented impacts of the pandemic on other biodiversity-themed citizen science project (Hochachka et al., 2021). Besides, there were large portions of observations with obscured geographic coordinates starting from July 2020 through end of the year. It may be that many observations were conducted at or near where the participant's home location and therefore observations were obscured to protect geoprivacy (i.e. preventing the data from revealing where the contributor lives) (iNaturalist, 2022a).

The geovisualization framework is meant to be an exploratory data analysis tool for examining contribution patterns (and pattern changes) in a VGI community. Empirically testing or verifying the hypotheses on what may have shaped the patterns (or what may have caused pattern changes) and explicating their implications can be left to subsequent formal data analyses.

4.3. Metrics for measuring participation

While conducting the above geovisual exploratory data analyses using the geovisualization tool (Sections 4.1 and 4.2), the analyst was constantly extracting the spatial, temporal, thematic, and social patterns presented by the visualizations and interpreting the similarities and differences between patterns across contributors. Understandably, these processes do not rely solely on the visualizations the geovisualization tool offers because the analyst's background knowledge also plays important roles in such processes. Nonetheless, the geovisualization tool provides visual aids to facilitate such endeavors. Specifically, the analyst noticed that the tool is instructive for evaluating the utility of existing participation measurement metrics and can inform developing additional metrics to depict participation characteristics along the four dimensions (i.e. spatial, temporal, thematic, and social). As an example, when visually comparing the spatial distribution of observations contributed by different contributors on the point location map, the analyst observed much variability in the spatial spread of observations across contributors. The standard distance across observation locations then came up naturally into the author's mind as a spatial metric reflecting such variability. As a result, standard distance was included (as one of the six features) for clustering contributor participation patterns (Section 3.1).

4.3.1. Spatial metrics

Several metrics have been proposed to measure the spatial characteristics of a VGI participant's data contributions: active area size, number of recording areas, and spatial aggregation (August et al., 2020). These metrics should be helpful for differentiating the spatial characteristics of participation patterns in iNaturalist, as the participation profiles explored in Section 4.1 (Figure 7) seem to have very different values on the three metrics.

New metrics can be developed to further distinguish a participant's spatial characteristics of participation (e.g. mobility and preferred observation environments). For instance, when comparing the spatial characteristics of contributions from different contributors using the point map and the pie chart showing observation composition by country, one may realize that the *number of visited countries* can reflect differential ability of the contributors to travel across country borders (e.g. the four contributors in Figure 7 visited 4, 2, 1 and 1 countries, respectively). Moreover, *land cover type(s) of the majority of observation locations* can indicate the participant's preferred observing landscape (in Figure 7, the contributors' major observation land cover types were deciduous broadleaf forests, deciduous broadleaf forests, grasslands, and urban and built-up lands, respectively). Value changes in such metrics can also capture pattern change over time. For example, due to travel constraints brought by the COVID-9 pandemic (Kishimoto & Kobori, 2021), a participant used to carry out observations in natural environments at international destinations may be forced to stay in the home country and contribute a larger share of observations in urban environments (Hochachka et al., 2021). As an example, the contributor depicted in Figure 9 traveled to another three countries outside of the United States in 2019 but stayed in the United States in 2020. Additionally, observations of a contributor form spatiotemporal trajectories reflecting the contributor's mobility. Statistics summarizing spatial characteristics of mobility trajectories, such as displacement, duration, speed, interval, radius and entropy (Wang et al., 2019), can also be considered as new spatial metrics for measuring participation.

4.3.2. Temporal metrics

Based on the conceptual framework of user-engagement (O'Brien & Toms, 2008), metrics have been developed to measure the temporal characteristics of participation (i.e. activity ratio, daily devoted time, weekly activity, relative activity duration, periodicity, and variation in periodicity) (Aristeidou et al., 2017; August et al., 2020; Boakes et al., 2016; Ponciano & Brasileiro, 2014). *Activity ratio*, *weekly activity*, *periodicity*, and *variation in periodicity* can be computed with information that is often available in VGI datasets to measure the overall activeness of a contributor and the temporal variability in the contributions. However, information (e.g. duration of observation session) is usually not available for computing *daily devoted time*. Also, *relative activity duration* (the ratio of days during which a participant remains linked to the project relative to the total number of days elapsed since the participant joined the project until the project is over) does not apply to most VGI projects because they are often open-ended and still continuously running (e.g. iNaturalist).

These metrics are derived simply based on whether the participant is actively contributing data on a particular day. They do not represent temporal variations in other aspects of contributor behavior or in data content. New metrics such as the median and interquartile range of number of species observations and identifications, observations to identifications ratio, and percentages of obscured observations per week (or per month) can be computed to indicate the temporal variabilities in data content, participant's contributing role tendency (observer vs. identifier), and geoprivacy awareness, respectively. Such metrics help reveal differences in the participation patterns along the temporal dimension across iNaturalist participants (e.g. the four "typical" contributors in Figure 7) or across time periods (e.g. the contributor in Figure 9).

4.3.3. Thematic metrics

Metrics were designed to reflect the thematic characteristics of data contents contributed by individual participants in biodiversity citizen science projects, such as proportion of taxa recorded, rarity recording, and single-species lists (see August et al., 2020 for definition and calculation). *Proportion of taxa recorded* and *rarity recording* can be computed on iNaturalist data to indicate whether a participant often records a few or many species (e.g. the four contributors in Figure 7) and to which extent the participant focuses on reporting only rare species. *Single-species lists* (the proportion of visits on which the participant submitted a single record) is not applicable to iNaturalist as records therein are individual species observations without indicating which observations resulted from the same "visit." This metric, however, may be useful for other VGI datasets, such as eBird where such information is available (Sullivan et al., 2009).

Additional metrics might be derived to reflect whether a participant tends to be a generalist (e.g. general naturalists) or a specialist (e.g. birders) from the perspective of certain taxonomic rank. For instance, the number of kingdoms covered by species observations contributed by a participant could indicate the span of the observed species biodiversity at the kingdom level (e.g. kingdom-level distributions Figure 7). One step further, the entropy of the kingdom-level species frequency distribution can be calculated to quantify if a participant spreads observation efforts across multiple kingdoms relatively evenly (i.e. higher entropy), or narrowly focuses on a few kingdoms (i.e. lower entropy).

4.3.4. Social metrics

Very few metrics have been developed to represent participation characteristics along the social dimension, although social interactions is the implicit backbone of many VGI communities (Sbrocchi et al., 2022). This might be attributed to the fact that many published VGI datasets do not contain necessary information for reconstructing community social interactions.

Social interactions in the iNaturalist project take the form of inter-participant species identifications among others. A social network representing such interactions can be reconstructed based on information contained in the "recordedby" (observer) and "identifiedby" (identifier) fields of the research-grade iNaturalist observations (Ueda, 2022). Nodes in the social network correspond to individual contributors and edges represent species identification interactions among contributors. Node-level metrics (e.g. centrality measures such as node degree and PageRank score) thus can be derived through social network analysis (McCulloh et al., 2013) to capture individual's participation characteristics along the social dimension (e.g. contributors in Figure 7). These metrics quantify a contributor's participation characteristics in the social network and differentiate participants with distinct social interaction patterns.

5. Discussion

5.1. Practical implications for iNaturalist

The patterns of iNaturalist participants' contributions uncovered through geovisual explorations in this study have practical implications for the iNaturalist project in terms of devising strategies for increasing participation or improving data quality. This study confirms participation inequality in iNaturalist (Table 2), a phenomenon common across many online communities that a small portion of contributors are highly active in contributing data whilst most others are ephemeral (Carron-Arthur et al., 2014). Highly active participants (e.g. cluster 4 contributors in Figure 7) are already very much self-motivated and they tend to stay in the community and contribute heavily (Zhang, 2020). To increase the overall participation, iNaturalist should elicit more contributions from low active or ephemeral contributors (cluster 1 contributors in Figure 7) by encouraging them to achieve higher levels of contribution (Haklay, 2016) (becoming clusters 2, 3, or 4 contributors by submitting more observations on a wider range of species and in many different types of habitat, being active across more months, sampling a larger geographic area, or identifying more of other contributors' observations). For example, iNaturalist could set up a system of contribution milestones as challenges for individual contributors, send out monthly e-mails to remind them of the gaps toward achieving the next milestone, and issue official recognitions or rewards (e.g. medal markers on profile page) when they achieve milestones. This same mechanism may also help alleviate contribution decline (Table 3) by keeping existing active contributors engaged. Moreover, the participation profiles of individual contributors (e.g. Figure 7) reflect contributor expertise and can be exploited for improving species identification accuracy in iNaturalist data. For instance, when resolving conflicting species identifications proposed by multiple contributors, the platform can put more weight on the opinions of contributors who had more experiences in observing and/or identifying related species (e.g. it is reasonable to assume that a contributor who has observed and/or identified a large number of fungi species before is more capable of identifying a given fungus observation).

5.2. Generalizability of the geovisualization framework

As demonstrated in Section 4, the proposed geovisualization framework is useful for conducting exploratory VGI data analyses that aim to examine individual VGI contributor's participation characteristics, detect participation pattern change and form hypotheses for explaining the change, and assess and develop metrics for measuring participation.

Conceptually, the geovisualization framework is expected to transfer to many VGI projects beyond iNaturalist as long as the VGI datasets contain information regarding the spatial, temporal, thematic, and social dimensions of participation. In principle, any VGI datasets at the minimum are spatially and temporally referenced with geographic coordinates and timestamps, explicitly or implicitly labeled with thematic information, and associated to individual contributors. Intervolunteer social interactions, though inherent to many VGI communities, may or may not be reflected in the resulted VGI datasets. For example, bird watchers maintain a large and active eBird community discussion group on Facebook to facilitate discussion among birders (Facebook, 2022), but social interactions therein are not integrated into the published eBird datasets (eBird, 2021). Nonetheless, even in cases where social interactions among VGI contributors cannot be reconstructed due to data unavailability, the proposed framework can still be utilized to facilitate volunteer's participation characteristics exploration, participation pattern change detection, hypothesis formation, metrics design, etc. along the spatial, temporal, and thematic dimensions.

From a technical point of view, the high-level architecture design of the geovisualization framework and the component software (e.g. PostgreSQL/PostGIS) and web programming libraries used to materialize the architecture design is expected to be generally applicable to any VGI datasets. Implementation specifics, however, may differ across datasets as geovisualization interfaces should be contextualized and tailored to the specific dataset and analytical problems at hand (Robinson, 2017). For example, different VGI datasets may well result in different database structures and therefore different SQL queries for retrieving relevant information. The specific visualizations may also be different because the most appropriate and effective way of visualizing data highly depends on dataset characteristics (data content, data format, etc.). Nevertheless, the architecture design of the geovisualization tool and its implementation for the iNaturalist dataset should offer a framework and backbone for others to build geovisual analytics tools customized to their respective VGI datasets of interest.

Lastly, the developed geovisualization framework facilitates exploratory VGI data analysis, which is often only one of the initial steps in a full-stack data analytics workflow. One should not expect to conduct end-to-end data analytics solely using the tool. Instead, this tool should be utilized in combination with other data analysis tools external to the system (e.g. statistical analysis, machine learning, etc.) in conducting comprehensive analyses of VGI participation characteristics. For instance, if the end goal is to understand iNaturalist contributors' participation patterns, this tool can help gain intuitions on the number of distinct clusters and the characteristics of each cluster. It can also assist developing new metrics to measure individual contributor's participation, which serve as input features to advanced analytical algorithms for clustering the contributors based on their participation patterns or for detecting pattern change over time. Moreover, based on intuitions gained on the contribution pattern changes (from 2019 to 2020) uncovered through geovisual exploratory data analyses in this study, one can conduct a subsequent study to investigate how the COVID-19 pandemic may have played a role in causing such changes by analyzing the association between patterns in the iNaturalist dataset and patterns in COVID-19 datasets.

5.3. Comparison with geovisualization tools for OpenStreetmap

A variety of geovisualization tools have been developed for analyzing data from OSM, which arguably is the most prominent VGI project in the world. Most OSM tools focus on the thematic contents of OSM data without paying attention to the underlying contributors. For instance, *OSM Inspector* (https://tools.geofabrik.de/

osmi/) is for users to visually inspect the geometries of OSM data to help with identifying and fixing geometric errors in the data (e.g. duplicate nodes, self-intersecting ways). *OSM History eXplorer* (https://hex.ohsome.org/) allows examining the spatial distribution and temporal trend of map features relating to different topics (health, building, land use, disaster, etc.). *OpenStreetMap Analytics* (http://osm-analytics.org/) visualizes the density of different types of OSM data (i.e. buildings, roads, rivers, amenities, places, and hospitals), provides a temporal histogram to show edits frequency over time, and offers a map for detecting gaps in OSM buildings data.

Unlike the above tools which are not capable of revealing OSM contributors' contribution patterns, several among the suite of OSM tools developed by Pascal Neis (https://resultmaps.neis-one.org/) could be useful in this regard. *OSMstats* (https://osmstats.neis-one.org/), has platform-level time series at various granularities (by year, by month, or by day) to show the temporal variations in contributors (e.g. active members, newly registered members), edits on map elements (created, modified, or deleted) and submitted changesets. At the individual level, OSMstats provides summary statistics (e.g. map changes, activity days) by individual contributors. *Who's around me?* (https://resultmaps.neis-one.org/oooc) maps out the location of individual contributors (e.g. main activity center). *Your OSM Heat Map* (https://yosmhm.neis-one.org/) visualizes editing hotspots for a given individual contributor. In addition, *OpenStreetMap Crowd* (Quinn & MacEachren, 2018; https://sterlingquinn.github.io/apps/crowdlens/) provides individual contributor-level visualizations based on subsets of OSM data in selected cities around the world. The visualizations can be altered by temporal and other filters. It is intended to help professional users of OSM to better understand the contribution characteristics of induvial OSM contributors in a specific place (i.e. city).

Compared to these OSM tools with focus on platform-level trends or on specific aspects of individual-level contribution, the geovisualization tool proposed in this study is tailored for visualizing and analyzing individual-level contribution patterns in iNaturalist while considering all of the four general dimensions of VGI contribution (i.e. spatial, temporal, thematic, and social). A full-scale comparison of contribution patterns between iNaturalist and OSM is beyond the scope of this study. A future study may utilize these tools to investigate the similarities and differences in contribution patterns among contributors of the two VGI projects.

6. Conclusion

This article describes a custom web-based geovisualization framework (as an exploratory data analysis tool) for exploring participation characteristics of individual VGI contributors. The framework provides a variety of interactive visualizations to aid human analysts to extract the spatial, temporal, thematic, and social interaction patterns in participation. I demonstrated the usefulness of the geovisualization framework through examples of use on iNaturalist data in three application scenarios: 1) gaining intuitions on the clustering and variabilities of individual participation patterns, 2) detecting participation pattern shifts over time and forming explanation hypotheses, and 3) assessing and developing metrics to measure participation. The geovisualization framework is expected to be applicable for exploring individual-level participation characteristics in VGI communities beyond iNaturalist. The author's experiences of utilizing the geovisualization framework for exploring the iNaturalist data was overall smooth and positive, although a comprehensive usability testing by other target users of the geovisualization framework was not conducted in this study. This work is among the first efforts to explore individual-level VGI participation characteristics using custom geovisualization and geovisual analytics tools. Such geovisual explorations can inform in-depth investigation of VGI contributors' participation characteristics through quantitative analysis in future work (e.g. testing hypotheses on what may have caused contribution pattern changes, and empirically evaluating the efficacy of the new metrics).

Acknowledgments

The author is deeply grateful to the numerous participants of the iNaturalist biodiversity citizen science project for their generous data contributions, which made this research possible. The constructive comments and suggestions from the anonymous referees, which helped improve the manuscript, are also greatly appreciated.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by the University of Denver through the Professional Research Opportunities for Faculty (PROF) Grant.

ORCID

Guiming Zhang http://orcid.org/0000-0001-7064-2138

Data availability statement

Data used in this research was downloaded from the iNaturalist website at https://www.inaturalist.org/observations/export and from the Global Biodiversity Information Facility website at https://www.gbif.org/dataset/50c9509d-22c7-4a22-a47d-8c48425ef4a7. The geovisualization framework (with data on a small number of selected iNaturalist contributors extracted from the full database) is accessible on GitHub at https://guiming.github.io/GeovizVGI/.

References

Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, A., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science*, *21*(8), 839–857. https://doi.org/10.1080/13658810701349011

Aristeidou, M., Herodotou, C., Ballard, H. L., Higgins, L., Johnson, R. F., Miller, A. E., Young, A. N., & Robinson, L. D. (2021). How do young community and citizen science Volunteers support scientific research on biodiversity? The case of iNaturalist. *Diversity*, *13*(7), 318. https://doi.org/10.3390/d13070318

Aristeidou, M., Herodotou, C., Ballard, H. L., Young, A. N., Miller, A. E., Higgins, L., Johnson, R. F., & Lepczyk, C. A. (2021, January). Exploring the participation of young citizen scientists in scientific research: The case of iNaturalist. *PLoS ONE*, *16*(1), 1–14. https://doi.org/10.1371/journal.pone.0245682

Aristeidou, M., Scanlon, E., & Sharples, M. (2017). Profiles of engagement in online communities of citizen science participation. *Computers in Human Behavior*, *74*, 246–256. https://doi.org/10.1016/j.chb.2017.04.044

Arleo, A., Didimo, W., Liotta, G., Miksch, S., & Montecchiani, F. (2022). Influence maximization with visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, *28*(10), 3428–3440. https://doi.org/10.1109/TVCG.2022.3190623

August, T., Fox, R., Roy, D. B., & Pocock, M. J. O. (2020). Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-67658-3

Bégin, D., Devillers, R., & Roche, S. (2018). The life cycle of contributors in collaborative online communities - the case of OpenStreetMap. *International Journal of Geographical Information Science*, *32*(8), 1611–1630. https://doi.org/10.1080/13658816.2018.1458312

Boakes, E. H., Gliozzo, G., Seymour, V., Harvey, M., Smith, C., Roy, D. B., & Haklay, M. (2016). Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. *Scientific Reports*, *6*(1), 1–11. https://doi.org/10.1038/srep33051

Carron-Arthur, B., Cunningham, J. A., & Griffiths, K. M. (2014). Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 principle and Zipf's law. *Internet Interventions*, *1*(4), 165–168. https://doi.org/10.1016/j.invent.2014.09.003

Chen, J., MacEachren, A. M., & Guo, D. (2008). Supporting the process of exploring and interpreting space-time multivariate patterns: The visual inquiry toolkit. *Cartography and Geographic Information Science*, *35*(1), 33–50. https://doi.org/10.1559/152304008783475689

Chen, J., Roth, R. E., Naito, A. T., Lengerich, E. J., & MacEachren, A. M. (2008). Geovisual analytics to enhance spatial scan statistic interpretation: An analysis of U.S. cervical cancer mortality. *International Journal of Health Geographics*, *7*(1), 1–18. https://doi.org/10.1186/1476-072X-7-57

Connors, J. P., Lei, S., & 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. https://doi.org/10.1080/00045608.2011.627058

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. https://doi.org/10.3390/ijgi10070425

DiCecco, 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. https://doi.org/10.1093/biosci/biab093

eBird. (2021). *eBird basic dataset metadata (v1.13)*.

Enguehard, R. A., Hoeber, O., & Devillers, R. (2013). Interactive exploration of movement data: A case study of geovisual analytics for fishing vessel analysis. *Information Visualization*, *12*(1), 65–84. https://doi.org/10.1177/1473871612456121

Facebook. (2022). *eBird community discussion group*. https://www.facebook.com/groups/288737854555183/?mibextid=6NoCDW

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), 1–16. https://doi.org/10.1002/eap.2056

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

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

Guo, D., & Jin, H. (2011). Iredistrict: Geovisual analytics for redistricting optimization. *Journal of Visual Languages and Computing*, *22*(4), 279–289. https://doi.org/10.1016/j.jvlc.2011.03.001

Haklay, M. (2016). Why is participation inequality important? In C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, & R. Purves (Eds.), *European handbook of crowdsourced geographic information* (pp. 35–44). Ubiquity Press. https://doi.org/10.5334/bax

Haklay, M., & Weber, P. (2008). OpenStreetMap: User-generated street maps. *Pervasive Computing, IEEE*, *7*(4), 12–18. https://doi.org/10.1109/MPRV.2008.80

Hochachka, W. M., Alonso, H., Guti, C., Miller, E., & Johnston, A. (2021). Regional variation in the impacts of the COVID-19 pandemic on the quantity and quality of data collected by the project eBird. *Biological Conservation*, *254*, 108974. https://doi.org/10.1016/j.biocon.2021.108974

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), 1–21. https://doi.org/10.1371/journal.pone.0241957

iNaturalist. (2022a). *Frequently asked questions*. https://www.inaturalist.org/pages/help#geoprivacy

iNaturalist. (2022b). *iNaturalist observations*. https://www.inaturalist.org/observations

iNaturalist. (2022c). *iNaturalist observations export*. https://www.inaturalist.org/observations/export

Johnston, A., Hochachka, W. M., Strimas-Mackey, M. E., Ruiz Gutierrez, V., Robinson, O. J., Miller, E. T., Auer, T., Kelling, S. T., Fink, D., & Fourcade, Y. (2021). Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. *Diversity and Distributions*, *27*(7), 1265–1277. https://doi.org/10.1111/ddi.13271

Kelling, S., Lagoze, C., Wong, W.-K., Yu, J., Damoulas, T., Gerbracht, J., Fink, D., & Gomes, C. (2013). eBird: A human/Computer learning network to improve biodiversity Conservation and research. *AI Magazine*, *34*(1), 10–20. https://doi.org/10.1609/aimag.v34i1.2431

Kishimoto, K., & Kobori, H. (2021). COVID-19 pandemic drives changes in participation in citizen science project "city nature challenge" in Tokyo. *Biological Conservation*, *255*, 109001. https://doi.org/10.1016/j.biocon.2021.109001

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in K-Means clustering. *International Journal of Advance Research in Computer Science and Management Studies*, *1*(6), 90–95.

Kraak, M.-J. (2003). Geovisualization illustrated. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(5–6), 390–399. https://doi.org/10.1016/S0924-2716(02)00167-3

Kraak, M. J. (2008). From geovisualisation toward geovisual analytics. *Cartographic Journal*, *45*(3), 163–164. https://doi.org/10.1179/174327708X315183

Li, Z., Huang, X., Hu, T., Ning, H., Ye, X., Huang, B., Li, X., & Yang, C. (2021). ODT FLOW: A scalable platform for extracting, analyzing, and sharing multi-source multi-scale human mobility. *PLoS ONE*, *16*(8), e0255259. https://doi.org/10.1371/journal.pone.0255259

McCulloh, I., Armstrong, H., & Johnson, A. (2013). *Social network analysis with applications*. John Wiley & Sons.

Nakaya, T., & Yano, K. (2010). Visualising crime clusters in a space-time cube: An exploratory data-analysis approach using space-time kernel density estimation and scan statistics. *Transactions in GIS*, *14*(3), 223–239. https://doi.org/10.1111/j.1467-9671.2010.01194.x

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. https://doi.org/10.3390/ijgi1020146

Nelson, J. K., Quinn, S., Swedberg, B., Chu, W., & MacEachren, A. M. (2015). Geovisual analytics approach to exploring public political discourse on twitter. *ISPRS International Journal of Geo-Information*, *4*(1), 337–366. https://doi.org/10.3390/ijgi4010337

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955. https://doi.org/10.1002/asi.20801

Ponciano, L., & Brasileiro, F. (2014). Finding volunteers' engagement profiles in human computation for citizen science projects. *Human Computation*, *1*(2), 247–266. https://doi.org/10.15346/hc.v1i2.12

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. https://doi.org/10.3138/cart.51.2.3301

Quinn, S. D., & MacEachren, A. M. (2018). A geovisual analytics exploration of the OpenStreetMap crowd. *Cartography and Geographic Information Science*, *45*(2), 140–155. https://doi.org/10.1080/15230406.2016.1276479

Robinson, A. (2017). Geovisual analytics. In J. P. Wilson (Ed.), *The geographic information science & technology body of knowledge* (3rd quarte, Vol. 6). https://doi.org/10.22224/gistbok/2017.3

Robinson, A. C., Peuquet, D. J., Pezanowski, S., Hardisty, F. A., & Swedberg, B. (2017). Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data. *Cartography and Geographic Information Science*, *44*(3), 216–228. https://doi.org/10.1080/15230406.2016.1139467

Roth, R. E., & MacEachren, A. M. (2016). Geovisual analytics and the science of interaction: An empirical interaction study. *Cartography and Geographic Information Science*, *43*(1), 30–54. https://doi.org/10.1080/15230406.2015.1021714

Sarkar, D., & Anderson, J. T. (2022). Corporate editors in OpenStreetMap: Investigating co-editing patterns. *Transactions in GIS*, *26*(4), 1879–1897. https://doi.org/10.1111/tgis.12910

Sauermanna, H., & Franzonib, C. (2015). Crowd science user contribution patterns and their implications. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(3), 679–684. https://doi.org/10.1073/pnas.1408907112

Sbrocchi, C., Pecl, G., van Putten, I., & Roetman, P. (2022). A citizen science community of practice: Relational patterns contributing to shared practice. *Citizen Science: Theory and Practice*, *7*(1), 1–14. https://doi.org/10.5334/CSTP.358

Sulla-Menashe, D., & Friedl, M. A. (2018). *User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product*. https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf

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. https://doi.org/10.1016/j.biocon.2009.05.006

Tomaszewski, B., & MacEachren, A. M. (2012). Geovisual analytics to support crisis management: Information foraging for geo-historical context. *Information Visualization*, *11*(4), 339–359. https://doi.org/10.1177/1473871612456122

Tupikina, L., Schlosser, F., Voskresenskii, V., Kloppenborg, K., Lopez, F., Mariz, A., Mogilevskaja, A., Haklay, M., & Tzovaras, B. G. (2021). iNaturalist citizen science community during city nature challenge: New computational approach for analysis of user activity. *ArXiv Preprint ArXiv: 2112.02693*.

Ueda, K. (2022). *iNaturalist research-grade observations. iNaturalist.org. occurrence dataset*. https://doi.org/10.15468/ab3s5x accessed via GBIF.org

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. https://doi.org/10.1080/00219266.2020.1739114

Van Ho, Q., Lundblad, P., Aström, T., & Jern, M. (2012). A web-enabled visualization toolkit for geovisual analytics. *Information Visualization*, *11*(1), 22–42. https://doi.org/10.1177/1473871611425870

Wang, J., Kong, X., Xia, F., & Sun, L. (2019). Urban human mobility: Data-driven modeling and prediction. *ACM SIGKDD Explorations Newsletter*, *21*(1), 1–19. https://doi.org/10.1145/3331651.3331653

Wang, Z., & Ye, X. (2018). Social media analytics for natural disaster management. *International Journal of Geographical Information Science*, *32*(1), 49–72. https://doi.org/10.1080/13658816.2017.1367003

White, J. J., & Roth, R. E. (2010, September 15). TwitterHitter: Geovisual analytics for harvesting insight from volunteered geographic information. *Proceedings of GIScience* Zurich, Switzerland, (Vol. 2010).

Xin, Y. (2022). Anomaly detection for volunteered geographic information: A case study of safecast data. *International Journal of Geographical Information Science*, *36*(7), 1423–1442. https://doi.org/10.1080/13658816.2021.1981333

Yan, Y., Feng, C., Huang, W., Fan, H., Wang, Y., & 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. https://doi.org/10.1080/13658816.2020.1730848

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. https://doi.org/10.3390/ijgi9100597

Zhang, G. (2021). Volunteered geographic information. In J. P. Wilson (Ed.), *Geographic information science & technology body of knowledge* (1st quarte). https://doi.org/10.22224/gistbok/2021.1.1

Zhang, G. (2022). Detecting and visualizing observation hot-spots in massive volunteer-contributed geographic data across spatial scales using GPU-accelerated kernel density estimation. *ISPRS International Journal of Geo-Information*, *11*(1), 55. https://doi.org/10.3390/ijgi11010055

Zhang, T., Wang, J., Cui, C., Li, Y., He, W., Lu, Y., & Qiao, Q. (2019). Integrating geovisual analytics with machine learning for human mobility pattern discovery. *ISPRS International Journal of Geo-Information*, *8*(10), 434. https://doi.org/10.3390/ijgi8100434

Zhang, G., & Zhu, A.-X. (2018). The representativeness and spatial bias of volunteered geographic information: A review. *Annals of GIS*, *24*(3), 151–162. https://doi.org/10.1080/19475683.2018.1501607

Zhang, G., & Zhu, A.-X. (2019a). A representativeness directed approach to spatial bias mitigation in VGI for predictive mapping. *International Journal of Geographical Information Science*, *33*(9), 1873–1893. https://doi.org/10.1080/13658816.2019.1615071

Zhang, G., & Zhu, A.-X. (2019b). A representativeness heuristic for mitigating spatial bias in existing soil samples for digital soil mapping. *Geoderma*, *351*(May), 130–143. https://doi.org/10.1016/j.geoderma.2019.05.024

Zhang, G., Zhu, A.-X., Huang, Z.-P., Ren, G., Qin, C.-Z., & Xiao, W. (2018). Validity of historical volunteered geographic information: Evaluating citizen data for mapping historical geographic phenomena. *Transactions in GIS*, *22*(1), 149–164. https://doi.org/10.1111/tgis.12300

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. https://doi.org/10.1111/tgis.12303

Zook, M., Graham, M., Shelton, T., & Gorman, S. (2010). Volunteered geographic information and crowdsourcing disaster relief: A case study of the Haitian earthquake. *World Medical & Health Policy*, *2*(2), 6–32. https://doi.org/10.2202/1948-4682.1069