Defining Local Experts: Geographical Expertise as a Basis for Geographic Information Quality

Colin Robertson[1] and Rob Feick[2]

1 **Department of Geography and Environmental Studies, Wilfrid Laurier University, Waterloo, ON, Canada**
`crobertson@wlu.ca`

2 **School of Planning, University of Waterloo, Waterloo, ON, Canada**
`robert.feick@uwaterloo.ca`

Abstract

As more data are produced by location sensors, mobile devices, and online participatory processes, the field of GIScience has grappled with issues of information quality, context, and appropriate analytical approaches for data with heterogeneous and/or unknown provenance. Data quality has often been viewed through a bifurcated lens of experts and amateurs, but consideration of what the nature of geographical expertise is reveals a much more more nuanced situation. We consider how adapting frameworks from the field of studies of experience and expertise may provide a conceptual basis and methodological framework for evaluating the quality of geographic information. For contributed geographic information, quality is typically derived from a data user's trust in and/or perception of the reputation of the data producer. Trust and reputation of producers of geographic information has typically been derived from the presence or absence of professional qualifications and training. However this framework applies exclusively to 'crisp' notions of data quality, and has limited utility for more subjective contributions associated with volunteered geographic information which may provide a rich source of geographic information for many applications. We hypothesize that a conceptual framework for geographical expertise may be used as the basis for assessing information quality in both formal and informal sources of geospatial data. Two case studies are used to highlight the new concepts of geographical expertise introduced in the paper.

1998 ACM Subject Classification E.0 [Data] General

Keywords and phrases data quality, expertise, geographic information, conceptual framework

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.22

1 Introduction

What does it mean to be a 'local expert'? Local knowledge and expertise has long been a valued input into geographical inquiry, and geographic information systems and geoweb technologies are instrumental in acquiring, structuring, representing and disseminating georeferenced local knowledge in many application areas. The notion of local expertise rests on ideas of place [34], familiarity and personal experiences with locales [24]. These are distinct from and related to more formal types of geographical expertise (GE) such as metric or topological relations between and among places on the earth's surface. As geographical knowledge of all forms is increasingly codified into a variety of information products, a reconsideration of GE itself, what defines GE and how is it obtained, may provide a framework for evaluating and integrating heterogeneous sources of geospatial data.

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore; Article No. 22; pp. 22:1–22:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2 Related Work

2.1 Geographic Data Quality

Interest in approaches to conceptualize, measure, and improve spatial data quality has been a longstanding concern to GIScience research and practice. Devillers and Jeansolin [12] note that spatial data quality is often viewed from two interrelated perspectives - internal (producer) and external (user). Internal spatial data quality focuses on the characteristics of spatial data as a product and is concerned with measuring and documenting quality relative to known specifications that a dataset must satisfy. Specifications have a dual role since they provide data producers guidance for ensuring internal quality throughout data capture, compilation, and documentation procedures and they provide benchmarks for validating the quality of data products. The widely used ISO reference standard (ISO 19157:2013) decomposes internal data quality across six main dimensions (positional accuracy, thematic accuracy, completeness, logical consistency, temporal quality, usability) and outlines a corresponding set of quantifiable data quality measures.

External data quality, by contrast, focuses on the degree of correspondence between a user's needs in a given application environment and a dataset's internal characteristics [12]. Chrisman [5, 6] popularized the concept of viewing external quality in terms of 'fitness for use' to recognize that quality requirements vary among users and across different tasks. While intuitively appealing and conceptually elegant, broad-based implementation of the fitness for use concept has proven to be challenging in practice. User needs, even within a well-defined application context, are often diverse and dynamic as people gain a better understanding of the problem at hand or as decision making processes progress through successive stages (e.g. problem scoping, evaluation of options, etc. in land planning). Progress is being made within specific subfields and defined task arenas, however [27] note that additional research is required to develop domain ontologies and easy-to-use tools that enable users to define and assess fitness for use.

The past decade has seen our understanding of fitness for use and spatial data quality more generally challenged by increasing variety in what constitutes geographic data, who creates these data and how they are produced. The growth of personal computing and location aware devices has fuelled crowdsourcing, citizen science and volunteered geographic information (VGI) activities that have resulted in new sources and types of geographic data pertinent to many existing and novel applications. However, since the data sets that emerge from these processes are not products of a single entity with known and reliable data quality procedures (e.g. national mapping agency), traditional measures of data quality are not applicable. Heterogeneity is perhaps the common characteristics of these data sets. A single data set can contain wiki-like edits from many individuals who differ in terms of skills, motivations, and knowledge, data capture methods (e.g. screen digitizing, GPS), and geographic areas of interest [35]. Some types of VGI, such as OpenStreetMap, have a loosely-defined schema that provides a foundation for consistency that also allows for local adaptations [26], while others that are by-products of social processes (e.g. Twitter posts, geotagged photos, GPS trail mapping) are largely unstructured. In this context, quality measures such as accuracy, lineage and consistency can vary with contributor, area and on a feature-by-feature basis.

Since the heterogeneity within and across VGI and citizen science data sets precludes the use of formal internal data quality measures, quality assessment shifts from evaluating data products toward more inferred and producer-centric foci [32]. For example, the quality of OpenStreetMap data in a locale is often inferred from the number of local contributors or with reference to deviations in coverage or positional accuracy relative to expert-generated comparators [21, 15]. Many VGI sources centre on phenomena that are not collected by authoritative agencies (e.g. bird sightings) or contain more subjective observations (e.g. geotagged photos and comments about camping experiences) that cannot be ground truthed, but may be considered to potentially exhibit a degree of 'local expertise'. In both of these instances, the quality of data elements is often judged based on contributors' attributes such as formal qualifications, trustworthiness, reputation, or credibility [32, 28]. Bishr and Kuhn [1] refer to this transitivity between trust in an individual and trust in the spatial data they create as informational trust, which they demonstrate can vary spatially and temporally and Keßler and De Groot [23] extend to the feature level in OSM.

These more nuanced approaches to understanding the expertise of individuals (i.e. persons or single entities) to create spatial data that are fit for specific uses and locales have interesting parallels with what Golledge [18] referred to as the changing nature of geographic knowledge. The dominant change in geographic knowledge [18] identified was "a change from inventory dominated activity to the creation of knowledge generated by emphasizing cognitive demands" such as processes of logical reasoning, deduction, and geographic association. Our current data-rich environment is seemingly more focused on inventorying than searching for understanding, in part because the former is technologically and socially accessible. To address current data quality challenges, Goodchild and Li [20] advocate for a knowledge-based approach where geographic concepts are used to assess data quality. Operationalizing this notion of geographic knowledge is challenging and, we suggest, requires consideration of individuals' differing expertise to expertise across a range of activities such as inventorying ubiquitous features and facts (e.g. trace building outline) or contributing more specialized and locale-specific information (e.g. document biodiversity in local wetland).

2.2 Studies of Experience and Expertise

Academic study of expertise and experience has been the traditional domain of the sociology of science. In Collins and Evans' seminal paper on SEE [11], the distinction between contributory and interactional expertise within a scientific field was introduced. Embedded in this juxtaposition are two elements of expertise in science: the knowledge and capability to make contributions to the field, and the ability to participate and interact with other actors in the field. Contributory expertise is generally acquired through formal training and education and working within the domain of interest, while interactional expertise is gained through socialization and exposure to tacit knowledge. The traditional view of expertise is a one-dimensional construct based on accumulation of ability and experience, which leads to phase or stage-based models of expertise development [13]. Critically, this expertise is typically only recognized if it is acquired through educational programs that initiate socialization and immersion into the society of experts (i.e., contributory expertise can only be recognized through interactional expertise).

Some recent work in citizen science has attempted to classify aspects of citizen science in such one-dimensional models. Haklay's [22] typology of citizen scientists is an recent example of this, presenting a typology of citizen science projects from passive crowdsourcing to collaborative science. Coleman's [8] typology focuses on individual motivations, extending from neophyte through to expert amateur. Goodchild [19] is a rare paper explicitly examining GE itself, yet fails to provide any framework or methodology for evaluating or characterizing it, while highlighting the critical need for theorizing GE in the geoweb and neogeography era.

The motivating rationale for many VGI and citizen science projects is that many individuals may possess expertise that could be valuable as input into pressing environmental and societal issues. Collins and Evans [11] describe Wynne's [37] study, which contrasted the expertise of sheep farmers without qualifications with government scientists in the aftermath of a nuclear contamination incident. The study showed that sheep farmers had specialized contributory expertise relevant to the ecology of sheep in the region. More recently, Maderson and Wynne-Jones [25] describe the role of beekeeper knowledge in understanding causes of and solutions to colony collapse disorder. In many cases, the criterion for expertise relevant to an environmental problem should be *experience* rather than normative qualifications.

Collins [9] derives, based on long-term studies into expertise, three dimensions of expertise: that which is attained through accumulation of experience and enable ability to contribute to a specific domain (i.e. the traditional form), the tacit knowledge of a domain that can only be acquired through socialization within that domain, and the 'esotericity', or the degree to which the domain is esoteric (e.g., gravitational physics vs. car driving). This results in a three dimensional model of expertise, termed the Expertise-Space-Diagram (ESD) based on its graphical representation, providing a conceptual tool to investigate expertise (Figure 1a). In this paper, we explore the ESD for deconstructing and analyzing GE in a variety of contexts and forms.

2.3 Towards Geographical Expertise

The nature of GE in the context of information quality has not been explored in great detail. The data quality literature in GIScience is dominated by a paradigm borrowed from transactional data architectures where data models are defined *a priori* and discrepancy metrics can be formulated easily. Unfortunately, this approach has limited utility for messy, heterogeneous information sources, where often the question being asked is 'what can I do with these data?' rather than 'does this data meet the requirements of this application?'.

As discussed above, expertise can be defined along three dimensions; contributory, interactional, and esotericity. The forms of expertise commonly represented in geographic information vary widely, yet these concepts have not been formally incorporated into approaches for evaluating information quality. Creators of geographic information may have any combination of levels of expertise as it relates to a given type of geographic information. GE that is place-based, contextual, and general in nature today tends to be derived from experience and less frequently from training in regional specializations in academic geography. Goodchild [19] argues that locale familiarity was a cornerstone of GE in the regional tradition, and this has been greatly democratized by increasing travel, allowing more people to become familiar with more places. Many studies that employ analysis of place-based social media data attempt to capture local expertise [3, 4]. As well, many forms of indigenous knowledge is place-based, grounded in narratives and experiences, and conferred through traditions and oral histories [31]. Locale-familiarity, or place-based expertise, is a dimension of GE which has to do with experience of a particular place. This type of naïve geographical knowledge tends to be approximate, more topological than metric, and often prone to biases [14]. From their inception, GIS has had a difficult time representing these forms of knowledge [18], despite concerted research efforts to develop their capabilities. Spatial operations designed for crisp data models also have difficulty handling fuzzier spatial information sources representing these forms of expertise and new 'platial' analogues is currently an active research area [29, 16].

A related, but distinct form of GE is that which relates to knowledge of geographical archetypes. With the shift to understanding of geographic processes rather than regions, geographers have developed expertise in the outcome of spatial processes and forms thereof. We refer to groupings of these geographical expressions as place-types. A component of GE that relates to knowledge of place-types is a core aspect of many academic geographers. For example, experts in grassland ecosystems or suburban sprawl, may not have any experience with a particular locale, but will be able to use knowledge of underlying processes to make expert judgements and understanding of places. Valuable geographic information could therefore be contributed by individuals with high locale-familiarity or high place-type knowledge.

Geographical expertise that is technical, pertaining to the tools and practices of producing accurate geographic data has, until recently, been almost exclusively held by professionals in surveying and mapping sciences. However, many forms of data production that used to require such technical expertise no longer do, as new technologies simplify many tasks of geographic data production [19]. As some forms of expertise are attained by more people via their use of simple GIS or web-based editing tools, according to the Collins [9] model of expertise described above, this represents a change in the type of expertise required for this task (via reduced 'esotericity'), from specialist knowledge to ubiquitous expertise. This is a critical contribution of the ESD model, in that expertise itself is deconstructed relative to its ubiquity, such that expertise moves from being only possible for a select few (i.e., those that participate in an esoteric domain) to almost everyone. This model of expertise underlies more inclusive science-society relationships in general, as exemplified by the transition in citizen science from citizens-as-sensors (i.e., data collectors) to higher levels of citizen participation in design, scoping, analysis, and interpretation [2].

As GE pertains to the processes and patterns at or near the earth's surface, we provide a conceptual model of GE based on a translation of the ESD model in Figure 1. The key difference here is replacing the notion of specialist tacit knowledge in ESD with locale-familiarity and place-type knowledge on the X-axis in our model of GE. The tacit knowledge in SEE 'can be acquired only by immersion in the society of those who already possess it'. Locale-familiarity/place-type tacit knowledge can be attained by immersion within the locale of interest or, through immersion and study of locales of a similar place-type. While interactional expertise with a community of experts maps directly to the model of SEE, a prolonged immersion within a geographic locale can alone provide sufficient experience to enable locale-familiarity knowledge. Note also that locale-familiarity knowledge can indeed always be enhanced through socialization within a community of locale-familiarity experts, according a locale-familiarity-type of interactional expertise. Many individuals will have both locale-familiarity and place-type knowledge, as place-type is the generalization of locale-familiarity expertise into archetypes. In short, place-type expertise emphasizes the common characteristics across distinct locales, while locale-familiarity emphasizes the local uniqueness of places. The Y-axis pertains to the traditional view of expertise, as ability to make contributions to the field and is typically attained and/or recognized through formal education and training and accomplishments such as publications, expert testimony, posts on editorial boards, professional reputation etc. Finally, the Z-dimensions we have renamed to 'thematic specificity' which is how general or specific a geographically-based topic or issue is. For example, an individual with knowledge of major landmarks in a city would have lower thematic specificity than one with deep understanding of its road network evolution or immigrant social services network. As with the ESD, both specialist and ubiquitous theme specificity constitute valid forms of GE.

In order to link GE at the individual level to a conceptual framework for evaluating information quality, we need to consider how expertise manifests in geographic information and data products. Single-producer data can be assessed relative to the GE of the contributor. Note that single-producer in this context includes entities like national mapping agencies

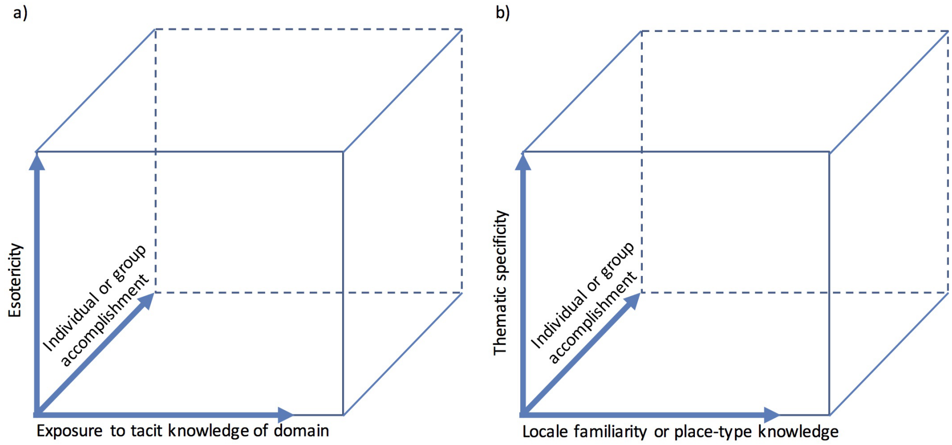

Figure 1 The dimensions of expertise according to the a) expertise-space diagram (ESD) and b) geographical expertise-space diagram (GESD).

that apply uniform data quality practices internally across their staff. Multi-authored data where contributors are at best loosely coordinated can be considered as a composite source of individual levels of GE.

Building on the Imitation Game methodology of SEE [10], if we consider two data products for an area such as two representations of a road centreline dataset in a city; one produced by professional mapping surveyors and another produced by several volunteers via OpenStreetMap. If both datasets are provided to quality control technicians for evaluation, and deemed to be of similar quality or potential utility, we can confer a level of expertise to the volunteer group based on their ability to produce a data product of equal value as traditional expert data. Such comparisons have been made in the literature many times, in efforts to demonstrate the quality of VGI or citizen-produced data. However the key point here is not the search for specific metric values but rather the interpretation of an information product by another member of the expert group. A related approach for local contribution data common of VGI, could be developed whereby community members judge each others' GE through collaborative experimental methods such as the imitation game.

3 A Framework for Geographical Expertise and Information Quality

Conceptualizing GE as a multi-dimensional surface provides a tool to position geographical information products according to the relative expertise of their authors in terms of their knowledge of locales, GI production, and the thematic specificity of the topic being represented. Such a framing enables deeper questions about information quality than what is typical for the fitness-for-use paradigm. As well, given the 'neogeographical' trend and shift towards heterogeneous data collectors and global scale data projects, explicit consideration of the components of GE articulates a more precise definition of quality in a given context, irrespective of potential application. The ESD for GI production is outlined in Figure 2, with the only difference from Figure 1 being the definition of Y-axis according to knowledge of geographical information production, ranging from those with no training and/or professional experience to GI professionals (e.g., surveyors, GI researchers).

To illustrate, four potential GI contributors are positioned within the GESD model, signified by letters. Person A might be a soil scientist collecting soil samples data in an

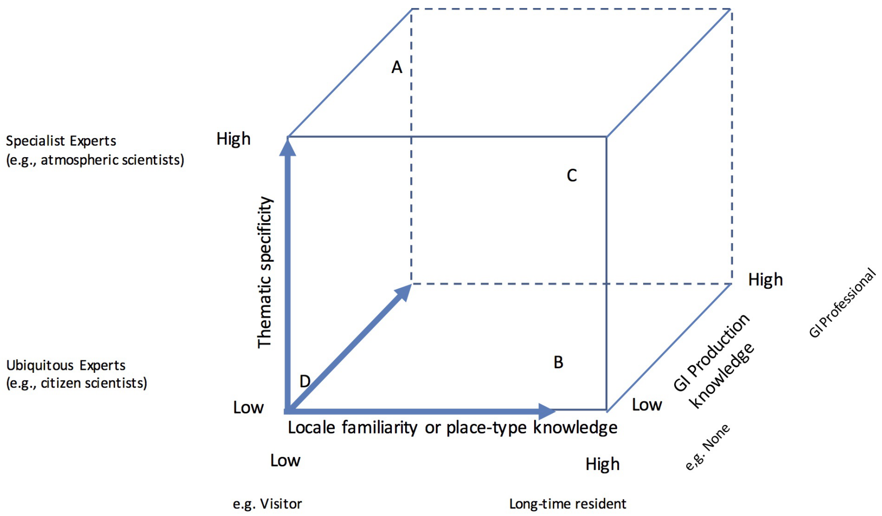

Figure 2 Geographical expertise diagram for geographical information production.

unfamiliar ecotype, exhibiting high thematic-specificity, low locale-familiarity/place-type, and high geographic-information-knowledge. Due to the high thematic-specificity of the GI production task, high geographic-information-knowledge is particularly important, whereas locale-familiarity/place-type is less so. Person B in Figure 2 could be a long-term resident inventorying heritage buildings without guidance, a task of low thematic-specificity, with high locale-familiarity/place-type and low geographic-information-knowledge associated with collecting georeferenced photographs of heritage buildings, basic data structuring, etc. Person C could be a local biologist creating a map of at-risk habitat based on a mix of input data, exhibiting high thematic-specificity, high locale-familiarity/place-type and relatively moderate geographic-information-knowledge. Such data may be of high quality depending on the technologies used for data collection. Finally, person D may be sharing personal photos or social media posts with vernacular place references or unconscious geotags; low in all dimensions, but still creating potentially useful GI.

Further examples of the GESD are provided in Figure 3 which compares two GI production tasks of different levels of thematic-specificity. Note that filled planes in Figures 3 and 4 refer to inclusion zones for producers of high quality GI, and dotted planes are for visual aid only. In landmark mapping, knowledge of city landmark locations is fairly ubiquitous, opening up the domain to participation by people with little training in geography. As well, the task of identifying features on a web-map or marking locations with a mobile app can be done fairly easily by individuals with little-to-no knowledge of GI production. As such, the plane of potential participants in terms of locale-familiarity/place-type and geographic-information-knowledge is large. Alternatively, a GI task higher in thematic-specificity such as developing a map of health risks due to air pollution would be possible for a much smaller subset of participants – those with higher geographic-information-knowledge and varying degrees of locale-familiarity/place-type. In this case, geographic-information-knowledge might be offset by their degree of local knowledge (e.g., experts in point-source interpolation may require less local knowledge). The plane of potential producers of high quality GI for this task is much smaller, and those falling outside of this plane might be more likely to produce erroneous and/or lower quality information products.

Thus far we have considered the GESD from the perspective of creators of GI. An alternate perspective is to frame minimal expertise required to contribute to existent GI sources within

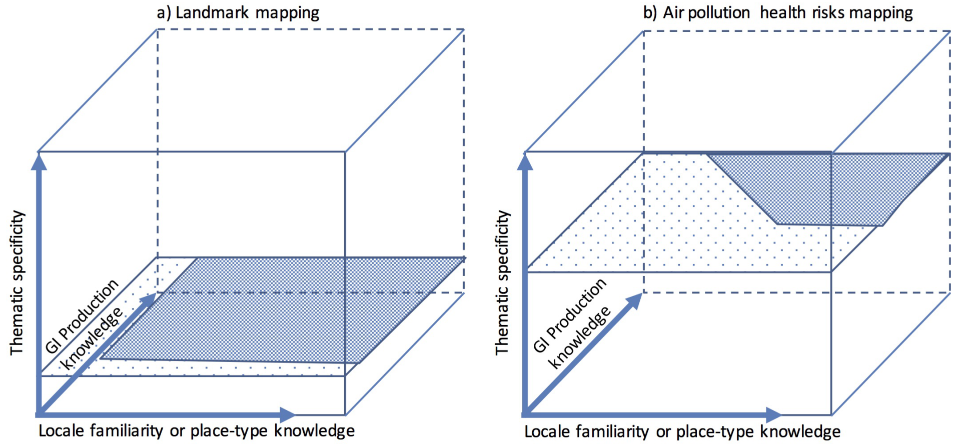

Figure 3 Geographical information production tasks in the geographical expertise diagram.

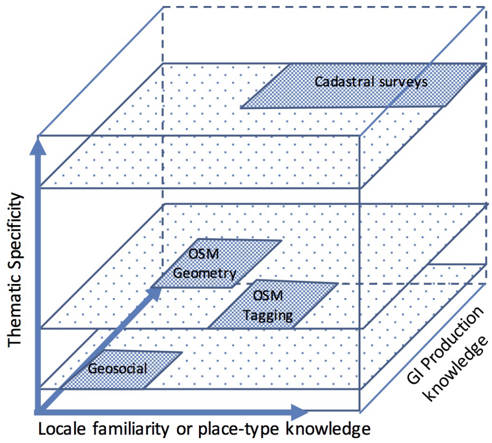

Figure 4 Geographical information products in the geographical expertise diagram.

the GESD, as is outlined in Figure 4. For cadastral surveys, a very high thematic-specificity application, high geographic-information-knowledge in the form of extensive training and professional experience, and a moderate level of locale-familiarity/place-type is required. VGI contributions to OSM constitute a more moderate thematic-specificity application, with different levels of GI and locale-familiarity/place-type required for tagging local features (i.e., attribution) and geometry editing. Geosocial data sources such as geotagged Twitter posts or Flickr photos require very little locale-familiarity/place-type, little geographic-information-knowledge, and are open to almost anyone with Internet access (i.e., low thematic-specificity).

While we have been considering locale-familiarity/place-type as a general category for inherent GE, there is an important scale dimension to all geographic knowledge. In general, geographers tend to develop expertise in one or several spatial scales, and GE as expressed in locale-familiarity/place-type is scale-dependent. We deconstruct spatial scale into two constituent components; grain and extent, and plot their relation to locale-familiarity/place-type in Figure 5. When spatial extent is very large and spatial grain is very small, the plane of potential expertise is limited, as heterogeneity dominates (i.e., one cannot be an expert in all individual areas over a large region). For small spatial extents, GE exists on the plane from

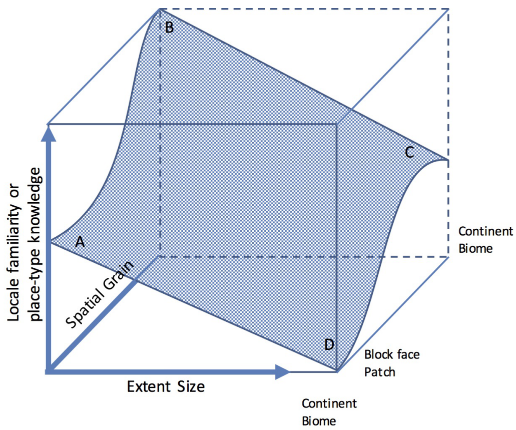

Figure 5 Spatial scale and geographical expertise.

moderate to high locale-familiarity/place-type as spatial grain gets larger (i.e., measurements are aggregated over large areas). For large extents, locale-familiarity/place-type varies from low to moderate as grain size increases.

Example individuals may further illustrate this relation between spatial scale and locale-familiarity/place-type illustrated in Figure 5. Individual A, with GE over a small study area, and a small unit size could represent a researcher studying how small irregularities in sidewalk conditions (e.g. cross-slope, missing and raised sidewalk sections) on a streets hinder mobility of persons using walking aids or wheelchairs. Individual B with GE for a small study area, large unit size might be an expert on history of urban development for a particular street in a major city. Individual C, with GE pertaining to a large study area and large grain size, such as global climate will have a limit on expertise when study area is huge, even if grain size is large, due to spatial heterogeneity. Finally, person D, were GE pertains to a large study area and small grain size, such as neighbourhood socioeconomic status in North American cities or wolf den habitat in the Boreal forest, has severe constraints on GE, as individuals cannot be an expert for all places over huge areas. Note also that the degree of spatial heterogeneity and/or autocorrelation impacts the potential plane for locale-familiarity and place-type.

4 Applied Examples

4.1 Geosocial Data Analysis – Mapping Emotional Affect

One of the ways that GE is implicitly considered in many studies that employ geosocial media data, is computationally distinguishing between 'locals' and 'tourists' [17]. Often this is done as a filtering step in data processing prior to more in-depth content or spatial analysis. The justification for this filtering stems from a desire to use these sources as a spatially explicit listening post in communities, to sample (albeit from a very unrepresentative sample frame) attitudes or activities in a community of interest. In our model of GE presented here, the filtering step would be estimating inclusion criterion for the geosocial potential plane mapped in Figure 4.

As part of an ongoing research project into geosocial media and urban stress, we investigated peoples' emotional affect at the time and place of contributing posts on Twitter. The rationale for this study was to validate sentiment analysis metrics for big data using

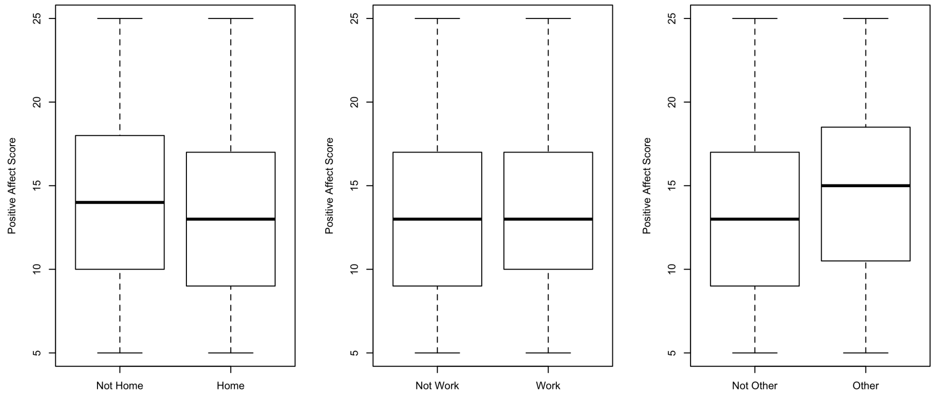

Figure 6 Positive affect scores for Twitter users at the time and place of Tweeting (significant differences in mean positive affect scores for home, not-home and work, not-work based on T tests, $\alpha = .05$).

in-situ psychometrically-valid survey questions [36]. Details of the study design are described elsewhere [33]. A total of 34 Twitter users participated in the study and contributed at least ten posts to their social media accounts during the study period selected here. Each user received short surveys upon entering and exiting the study, as well in response to their social media activity. In [30] we showed that participants were less likely to post messages that related to their immediate surroundings if they were at home or at work. Here we investigate whether participants emotional affect differed for participants based on Tweeting locale. Such information might be of interest for health planners hoping to gauge social media data for analysis of emotional expressions over space and time. From the perspective of GE, individuals' familiarity with the area would likely directly impact the degree to which mood is impacted by the environment.

We show in Figure 6 that positive affect scores differed when participants were Tweeting from home, work, or other locations. In terms of GE, we might be most interested in spatial patterns of emotional affect for residents (with greater experience and locale-familiarity) than with users Tweeting during work or leisure activities. Naïve mining of geosocial data streams may ignore these personal place-based variations in emotional affect.

4.2 Environmental Citizen Science – GrassLander Project

Data quality is a persistent issue in environmental citizen science research and practice, as assumptions about the expertise inherent in the categories of 'citizen' and 'scientist' dominate perceptions of information value. As Cinnamon [7] illustrates, such dichotomies are far from the norm in most VGI projects, where participants in projects cover a range of skill levels, experience, and training. We initiated a citizen science project eliciting input from farmers about agriculture activities and bird observations seen on their farms. This web-based application (www.grasslander.org) required farmers to go through a farm-set up phase in which they selected land parcel polygons for their farm boundaries and digitized their farm fields using web-based spatial data editing tools. We assume that farmers, whose livelihoods depend on their land, have high locale-familiarity/place-type, and as such can adequately provide geospatial data on their properties, regardless of their

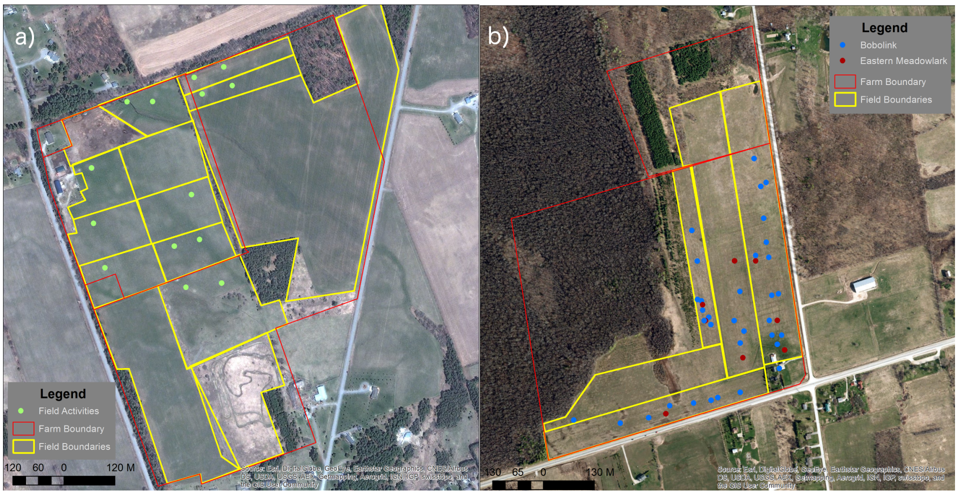

Figure 7 Geographical information contributed by 'citizen scientists' on farm geometry, a) field activities, and b) bird observations in southwest Ontario, Canada).

geographic-information-knowledge if provided with simple tools for creating geographic data. As well, farmers were asked to report on observations of grassland birds (bobolinks and eastern meadowlarks) sighted on their properties.

Geographical information produced by two participants are provided in Figure 7. In both examples, field geometry was found to coincide well with visible fields from aerial imagery. Digitized geometry evident from the participant in Figure 7a demonstrates high attention to spatial details, excluding between-field areas and careful digitizing around their home, however a section of field extends beyond the farm boundary. The participant data in Figure 7b shows field geometry that would not be known from existing aerial imagery. In the context of our GE model, the provision of intuitive web-based digitizing tools lowered the required geographic-information-knowledge needed for producing high quality data in this application. As well, the participant contributed many bird sighting records throughout their property. In future work, we aim to evaluate the degree of locale-familiarity vs. place-type expertise by having participants comment on and characterize farms other than their own, having individuals serve as 'experts' on their own local properties. Such a system can then be incorporated into statistical measures of GE using the imitation game methodology commonly deployed in SEE [10].

5 Conclusions

The model for GE presented here provides three dimensions of GE based on the ESD model of Collins [9]. We demonstrate how concepts of geographic knowledge can be embedded within the ESD model and concepts from SEE can be adapted to geographical information. Our framework provides the tools to deconstruct both data contributors and geographical information products along three dimensions of GE. Such a deconstruction may serve as a basis for developing information quality metrics robust to heterogeneity in the contributions common, but not limited to, many forms of volunteered and/or 'neo' geographic information.

Acknowledgements. The authors would like to thank the participants that contributed to both case studies presented in this research.

References

1 M. Bishr and W. Kuhn. Trust and reputation models for quality assessment of human sensor observations. In *International Conference on Spatial Information Theory*, pages 53–73. Springer, 2013.

2 Rick Bonney, Jennifer L. Shirk, Tina B. Phillips, Andrea Wiggins, Heidi L. Ballard, Abraham J. Miller-Rushing, and Julia K. Parrish. Next Steps for Citizen Science. *Science*, 343(6178):1436–1437, March 2014. doi:10.1126/science.1251554.

3 Z. Cheng, J. Caverlee, H. Barthwal, and V. Bachani. Finding Local Experts on Twitter. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, pages 241–242, New York, NY, USA, 2014. ACM.

4 Z. Cheng, J. Caverlee, H. Barthwal, and V. Bachani. Who is the Barbecue King of Texas?: A Geo-spatial Approach to Finding Local Experts on Twitter. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '14, pages 335–344, New York, NY, USA, 2014. ACM.

5 N. Chrisman. Part 2: Issues and Problems Relating to Cartographic Data Use, Exchange and Transfer: The Role Of Quality Information In The Long-Term Functioning Of A Geographic Information System. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 21(2-3):79–88, July 1984.

6 N. Chrisman. Development in the treatment of spatial data quality. *Fundamentals of Spatial Data Quality*, pages 21–30, 2006.

7 J. Cinnamon. Deconstructing the binaries of spatial data production: Towards hybridity. *The Canadian Geographer / Le Géographe canadien*, 59(1):35–51, March 2015.

8 D. Coleman, Y. Gerogiadou, and J. Labonte. Volunteered Geographic Information: the nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research*, 4:332–358, 2009.

9 H. Collins. Three dimensions of expertise. *Phenomenology and the Cognitive Sciences*, 12(2):253–273, April 2011.

10 H. Collins, R. Evans, M. Weinel, J. Lyttleton-Smith, A. Bartlett, and M. Hall. The Imitation Game and the Nature of Mixed Methods. *Journal of Mixed Methods Research*, page 1558689815619824, December 2015.

11 H. M. Collins and R. Evans. The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science*, 32(2):235–296, April 2002.

12 R. Devillers and R. Jeansoulin, editors. *Fundamentals of spatial data quality*. ISTE, 2006.

13 H. L. Dreyfus and S. E. Dreyfus. From Socrates to Expert Systems: The Limits of Calculative Rationality. In C. Mitcham and A. Huning, editors, *Philosophy and Technology II*, number 90 in Boston Studies in the Philosophy of Science, pages 111–130. Springer Netherlands, 1986.

14 M. J. Egenhofer and D. M. Mark. Naive Geography. In *Spatial Information Theory A Theoretical Basis for GIS*, pages 1–15. Springer, Berlin, Heidelberg, September 1995.

15 Hongchao Fan, Alexander Zipf, Qing Fu, and Pascal Neis. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*, 28(4):700–719, 2014.

16 S. Gao, K. Janowicz, and G. McKenzie. Towards Platial Joins and Buffers in Place-Based GIS. In *Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place*, COMP '13, pages 42:42–42:49, New York, NY, USA, 2013. ACM.

17 F. Girardin, F. Calabrese, F. D. Fiore, C. Ratti, and J. Blat. Digital Footprinting: Uncovering Tourists with User-Generated Content. *IEEE Pervasive Computing*, 7(4):36–43, October 2008.

18 R. G. Golledge. The Nature of Geographic Knowledge. *Annals of the Association of American Geographers*, 92(1):1–14, March 2002.

19 M. Goodchild. NeoGeography and the Nature of Geographic Expertise. *J. Locat. Based Serv.*, 3(2):82–96, June 2009.

20 M. F. Goodchild and L. Li. Assuring the quality of volunteered geographic information. *Spatial statistics*, 1:110–120, 2012.

21 M. Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4):682–703, 2010.

22 M. Haklay. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In Daniel Sui, Sarah Elwood, and Michael Goodchild, editors, *Crowdsourcing Geographic Knowledge*, pages 105–122. Springer Netherlands, January 2013.

23 Carsten Keßler and René Theodore Anton De Groot. Trust as a proxy measure for the quality of volunteered geographic information in the case of OpenStreetMap. In *Geographic information science at the heart of Europe*, pages 21–37. Springer, 2013.

24 M. Lewicka. Place attachment: How far have we come in the last 40 years? *Journal of Environmental Psychology*, 31(3):207–230, September 2011.

25 S. Maderson and S. Wynne-Jones. Beekeepers' knowledges and participation in pollinator conservation policy. *Journal of Rural Studies*, 45:88–98, June 2016.

26 S. Maguire and M. Tomko. Ripe for the picking? Dataset maturity assessment based on temporal dynamics of feature definitions. *International Journal of Geographical Information Science*, 0(0):1–25, February 2017.

27 I. Poças, J. Gonçalves, B. Marcos, J. Alonso, P. Castro, and J. P. Honrado. Evaluating the fitness for use of spatial data sets to promote quality in ecological assessment and monitoring. *International Journal of Geographical Information Science*, 28(11):2356–2371, 2014.

28 J. Preece. Citizen Science: New Research Challenges for Human–Computer Interaction. *International Journal of Human-Computer Interaction*, 32(8):585–612, 2016.

29 T. Quesnot and S. Roche. Platial or Locational Data? Toward the Characterization of Social Location Sharing. In *2015 48th Hawaii International Conference on System Sciences*, pages 1973–1982, January 2015.

30 C. Robertson, R. Feick, M. Sykora, K. Shankardass, and K. Shaughnessy. Personal Activity Centres and Geosocial Data Analysis: Combining big data with small data. In *AGILE 2017*, Wageningen, Netherlands, May 2017. Springer-Verlag.

31 R. A. Rundstrom. GIS, Indigenous Peoples, and Epistemological Diversity. *Cartography and Geographic Information Systems*, 22(1):45–57, January 1995. `doi:10.1559/152304095782540564`.

32 H. Senaratne, A. Mobasheri, A. L. Ali, C. Capineri, and M. Haklay. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1):139–167, 2017.

33 M. D. Sykora, C. Robertson, K. Shankardass, R. Feick, K. Shaughnessy, B. Coates, H. Lawrence, and T. Jackson. Stresscapes: validating linkages between place and stress expression on social media. In *Proceedings of the 2nd International Workshop on Mining Urban Data*, pages 80–84, Lille, France, 2015.

34 Y.-F. Tuan. Place: An Experiential Perspective. *Geographical Review*, 65(2):151–165, April 1975. ArticleType: research-article / Full publication date: Apr., 1975 / Copyright © 1975 American Geographical Society.

35 M. Van Exel, E. Dias, and S. Fruijtier. The impact of crowdsourcing on spatial data quality indicators. *Proceedings of GiScience 2011*, 2010.

36 D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070, 1988.

37 B. Wynne. Sheepfarming after Chernobyl: A Case Study in Communicating Scientific Information. *Environment: Science and Policy for Sustainable Development*, 31(2):10–39, March 1989.