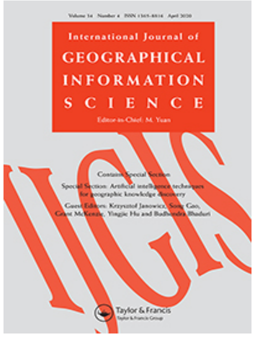

Taylor & Francis
Taylor & Francis Group

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond

Krzysztof Janowicz, Song Gao, Grant McKenzie, Yingjie Hu & Budhendra Bhaduri

To cite this article: Krzysztof Janowicz, Song Gao, Grant McKenzie, Yingjie Hu & Budhendra Bhaduri (2020) GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond, International Journal of Geographical Information Science, 34:4, 625-636, DOI: 10.1080/13658816.2019.1684500

To link to this article: https://doi.org/10.1080/13658816.2019.1684500

Published online: 30 Oct 2019.

Submit your article to this journal

Article views: 27155

View related articles

View Crossmark data

Citing articles: 96 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgis20

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
2020, VOL. 34, NO. 4, 625–636
https://doi.org/10.1080/13658816.2019.1684500

Taylor & Francis
Taylor & Francis Group

EDITORIAL

Check for updates

GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond

1. Introduction to GeoAI

Recent progress in Artificial Intelligence (AI) techniques, the large-scale availability of high-quality data, as well as advances in both hardware and software to efficiently process these data, are transforming a range of fields from computer vision and natural language processing to autonomous driving and healthcare. For instance, the availability of high-resolution geographic data and high-performance computing techniques together with deep learning fuel progress in fast and accurate object detection. Recent examples of GeoAI work include the detections of terrain features (Li and Hsu in this issue) and densely-distributed building footprints (Xie *et al*. in this issue), information extraction from scanned historical maps (Duan *et al*. in this issue), semantic classification (e.g. LiDAR point clouds) (Guo and Feng in this issue), novel methods for spatial interpolation (Zhu *et al*. in this issue), and advances in traffic forecasting (Polson and Sokolov 2017, Ren *et al*. in this issue). Similarly, machine learning and natural language processing are facilitating the extraction of geographic information from unstructured (textual) data, such as news articles and Wikipedia (Hu 2018) as well as the matching of natural features in multiple gazetteers (Acheson *et al*. in this issue).

At the same time, Semantic Web technologies, ontologies, and Linked Data are being employed to improve geographic information retrieval and to construct advanced geographic knowledge graphs for geo-enrichment (Ballatore *et al*. 2013, Regalia *et al*. 2019, Yan *et al*. 2019, Mai *et al*. 2019a) as well as semantically enabled services for spatial data infrastructures (Jones *et al*. 2014). A combination of multiple techniques aids in integrating autonomous vehicles with intelligent transport systems by incorporating real-time information gathered by traffic cameras and other sensors (Seif and Hu 2016, Zang *et al*. 2017). As these examples demonstrate, rapid progress is not isolated to specific downstream tasks or data types. Instead, we are observing how artificial intelligence techniques penetrate many aspects and activities across the sciences.

The use of AI techniques in geography and the earth sciences as such is not new. Openshaws' 1997 book on *Artificial Intelligence in Geography* serves as a prominent example (Openshaw and Openshaw 1997). Even before, Couclelis (Couclelis 1986) and Smith (Smith 1984) discussed the potential role of AI for geographic problem-solving in the 80s. What has changed since those early days cannot merely be attributed to novel computing architectures and advanced methods such as Generative Adversarial Networks (Goodfellow *et al*. 2014).

The current success of AI techniques is equally caused by a new culture of data creation and sharing. The exponential growth of data collected and curated over the past decade is not restricted to any specific type or medium but multi-modal and highly heterogeneous.

© 2019 Informa UK Limited, trading as Taylor & Francis Group

Social sensing, for instance, as the usage of data traces actively or passively emitted by humans via their near-body devices such as smartphones, would have been unthinkable in the 80s both due to technical limitations as well as for privacy concerns. While fields such as remote sensing have been (big) data-intensive since their early days, very high-resolution instruments do not merely provide more data, they dramatically grow the number of classes that can be distinguished, e.g. individual cars and debris. Similarly, the (open) availability of millions of near real-time car trajectories, images, reviews, recommendations, news, academic literature, and all kinds of sensor observations is a game-changer for those AI techniques that rely on large amounts of (high-quality) labeled data.

However, there may be something even more important than the pure availability of data and advanced methods combined, namely a change in culture. (I) As with open-source before, open-content makes data available to the masses. Individuals, institutions, and companies begin to realize that protecting their data in silos may be less beneficial than giving access to them. While data is rarely made available as a bulk download, companies now regularly share their (expensively acquired) data via APIs. This was simply unthinkable just two decades ago. *ProgrammableWeb*, for instance, reports more than 22,000 Web-APIs as of 2019 up from about 100 APIs in 2005. Put differently, the industry perceives the risk of somebody stealing their data as less impactful than their data remaining disconnected from the new data economy.

(II) Reusing data is the new normal. This may seem like a trivial point from today's perspective, but data reuse at scale is a new concept for many scientific domains. Designing an experiment, developing a conceptual model, deciding on measurement scales, selecting a sampling strategy, and collecting data are core parts of many scientific workflows. The opportunistic reuse and synthesis of data implies giving up control over some or even all of those steps. In contrast to (re)using imagery from long-running, well-known, quality-optimized, and technically well-understood earth observation satellites such as the Landsat series, reusing in-situ ecological or social science data is very different. Individuals have often collected these data with a particular setup and research question in mind. Typically, the accompanying metadata (if present at all) are not detailed enough to fully capture the contextual information required to understand whether a dataset is fit for the new purpose. In essence, all metadata records are incomplete as it is impossible to foresee future uses. The increase in research on data provenance and smart data workflows that automatically capture as much contextual information as possible is a response to this new reality (Gil *et al*. 2007, Moreau *et al*. 2008) in which data from all kinds of sources and domains are reused at scale.

(III) A new paradigm joins the empirical, theoretical, and computational paradigms that have characterized research before. This fourth paradigm (Hey *et al*. 2009) of data-intensive exploration highlights the increasing role of data synthesis (Janowicz *et al*. 2015) alongside analysis. This implies that one data source can be used as a proxy for another, more difficult to acquire dataset. It also suggests that combining multiple data sources may support a more holistic understanding of a research question or may help in mitigating problems of data sparsity or representational bias. To give a concrete example, Jacobs *et al*. (2009) used a readily available network of thousands of (low-quality) webcam to determine the onset of spring leaf growth. Gao *et al*. (2017) showed how social media from a variety of sources can be used to detect and delineate vague cognitive regions and how the extracted regions that resemble those acquired from direct human participants testing.

While none of these three identified aspects alone is necessarily new, the arising *data culture* certainly is. Consider, for instance, the following observation by Mike Goodchild: Most early research utilizing volunteered geographic information (VGI) did so to *confirm* or reproduce findings or theories that have been brought up before.[1] It is only in the last few years that VGI has been used to reveal *new* insights, question existing theories, or even propose new theories altogether. In fact, and in line with new research directions such as Web Science, researchers are beginning to study the ecosystem of geographic information as such, e.g. via geographic information observatories (Adams *et al*. 2014, Janowicz *et al*. 2014).

Summing up, GeoAI as a subfield of spatial data science utilizes advancements in techniques and data cultures to support the creation of more intelligent geographic information as well as methods, systems, and services for a variety of downstream tasks. These include image classification, object detection, scene segmentation, simulation and interpolation, link prediction, (natural language based) retrieval and question answering, on-the-fly data integration, geo-enrichment, and many others.

2. Spatially explicit models

Ideally, the application of techniques from artificial intelligence and data science to spatial data in the earth and social sciences is not a one-way street. Recent research (Yan *et al*. 2017, 2018, 2019, Chu *et al*. 2019, Mac Aodha *et al*. 2019) has shown that spatially explicit models substantially outperform more general models when applied to spatial data. Interestingly, designing neural architectures for spatially explicit models can also be regarded as introducing an inductive bias (Battaglia *et al*. 2018). However, what exactly are spatially explicit models and what do they have in common? How can we integrate spatial and temporal aspects to various machine learning-based techniques, and how much spatial data are required for these models to make a difference?

Interestingly, while there is no shortage of spatially explicit models and methods to address the needs of specific domains or applications, the question of what makes a model spatially explicit in the first place received less attention. Notable examples include the work of Goodchild and Janelle (2004) and Kuhn (2012). For instance, a model can be called spatially explicit if it fulfills the following requirements (Goodchild 2001):

- Invariance test: The results of spatially explicit models are not invariant under relocation of the studied phenomena.
- Representation test: spatially explicit models contain spatial representations of the studied phenomena in their implementations (this can be in the form of coordinates, spatial relations, place names, and so on).
- Formulation test: spatially explicit models make use of spatial concepts in their formulations, e.g. the notion of a neighborhood.
- Outcome test: The spatial structures/forms of inputs and outcomes of the model differ.

Spatially explicit models are those that satisfy at least one of these tests (and thereby any of their combinations). For instance, imagine a simple dataset that contains cities, their geographic location, as well as their population. A mere population-based ranking of those cities is not spatially explicit as their location (representation) is not part of the analysis. In contrast, answering whether densely populated cities are clustered would require a spatially explicit perspective. This, in turn, should not be confused with an analysis that may reveal spatial insights without being spatially explicit itself. Take, for instance, an alphabetic ordering of the cities (and places more generally) at California's coast. Such a list would reveal geographic insights about the origin of expeditions and the times they visited or established these places.

As is, however, the tests mentioned above cannot be carried out experimentally, nor can their degree or relevance be measured. More concretely, what is the tradeoff between designing a machine learning architecture that explicitly accounts for space versus a more general setup that would have to learn to value space implicitly? Will these more general models catch up given enough data without the need to increase the complexity of the architecture? What portion of a dataset has to be spatial in order to justify spatially explicit models? Those taking a strong position in favor of general models will have to justify why progress on neural architectures is required at all, if the availability of data is the only variable that matters. Similarly, those that favor domain-specific models will have to justify why developing more complex models is superior to providing more labeled data. Both these stances and any middle ground between them will have to address the question of how spatial (and place-based) aspects should be represented in data across domains and whether our current way of largely thinking in terms of fields and objects is still adequate at a time where graph data and the power to linking statements across domains are at the forefront. Finally, geographic identifiers often play a key role as a nexus that connects actors, events, and objects together across and within data hubs, e.g. on the global Linked Data cloud. Hence, spatio-temporally scoping data (Silva *et al*. 2006, Adams *et al*. 2015) will only increase this trend and make spatial aspects part of many everyday information retrieval tasks such as the semantic annotation of news.

Put differently, successful GeoAI research will have to address why (geo-)spatial matters by making a case for spatially explicit models. It will also have to showcase how graph data and new methods developed on the symbolic and sub-symbolic levels can easily be integrated into today's GIS workflows (Mai *et al*. 2019a).

3. Question answering and summarization

GeoAI research will also contribute to question answering and smart digital assistants more generally. This follows from the increasing availability and importance of spatial data such as place names and spatial relations discussed before, but also from the fact that digital assistants are quickly becoming part of our everyday lives. This gives these systems access to a plethora of contextual information and enables them to answer more personalized questions. For instance, instead of asking about the construction date of the Eiffel Tower or how long it will take to drive to the airport, future users may ask for vacation locations their parents would have visited, an audiobook about the region they are currently driving through, or simply a central but quiet hotel. These and similar questions require an additional step, namely identifying a user's location, distances to other features, reasoning about topological relations, understanding vague cognitive regions, and so on. Hence, current approaches, e.g. those directly utilizing sentence embedding models (Arora *et al*. 2017) or other forms of computing text similarity, may fall short.

While the mode in which questions are asked is relevant for the selection of appropriate methods, many higher-level challenges remain the same. For instance, how to summarize geographic information while answering more open-ended questions such as for important facts about Los Angeles (Yan *et al*. 2019). From a knowledge graph perspective, there are tens of thousands of statements (e.g. triples) about every major city. So what is special about one city in particular? An answer to this cannot merely be technical; it has to address the question of what makes a good and fair summary in the first place. For instance, and inspired by Rodriguez and Egenhofer (2004), one could argue that a summary has to account for both *commonality* and *variability*. This is particularly important for comparative questions or those that involve heterogeneous regions. To give a concrete example, an answer to the question about Los Angeles may state that while it is similar to the nearby city of San Diego in terms of climate, beaches, belonging to California, and so on, it is unique due to its motion picture industry.

One can also start with the answer and study how to discover and share GIS functionality based on the questions they are designed to answer (Scheider *et al*. 2019). Similarly, one may rethink the entire interaction with modern GIS and abstract it to a higher level centered around the scientific questions to be answered instead of the technical steps involved in doing so (Vahedi *et al*. 2016). Finally, one can study how to relax questions to arrive at an approximate (or at least related) answer (Wang *et al*. 2018, Mai *et al*. 2019b). There are many reasons for doing so. For example, when the initial question cannot be answered due to the sparsity of the knowledge graph or when a user is not sufficiently familiar with the ontology used to represent data.

4. Social sensing

Machine learning and artificial intelligence methods also have an important role to play in what is often referred to as *social sensing* (Aggarwal and Abdelzaher 2013, Liu *et al*. 2015, Janowicz *et al*. 2019). It can be defined as the use of (user-generated) digital content to better understand human dynamics. Social sensing has been applied to a range of tasks from identifying human mobility patterns (Li *et al*. 2019) and exploring structure in social networks, to urban planning solutions (Zheng *et al*. 2014, Resch *et al*. 2015, Zeile and Resch 2018) with varying degrees of success. The process of social sensing involves the creation of *semantic signatures* (Janowicz *et al*. 2019), multi-dimensional data signatures (i.e. spatial, temporal, and thematic features) that are extracted from the digital trace that is left behind as people's digital lives interact with their physical activities. This digital trace is increasingly produced and collected through sensor-rich mobile and IoT devices. The plethora of sensors available on today's mobile device means that the data being produced not only includes information pertaining to one's location, but also attributes such as the ambient temperature, luminosity, noise level, and so on. The sheer amount of data collected via these devices, as well as the heterogeneity of the actual content, make these data particularly well suited to the analysis through novel techniques situated in GeoAI (Martin *et al*. 2018). Social sensing and semantic signatures have contributed to fields such as public health (Chaix 2018), activity prediction (Regalia *et al*. 2016), and privacy preservation (Khan *et al*. 2019), to name a few.

In much of the social sensing research, semantic signatures are often rooted in the concept of *place*, using place as the reference system through which to compare different activities, dynamics, and social interactions. The digital trace collected through mobile sensors mentioned previously is often referred to as passive data collection with context being inferred from passive sensors. User-contributed data such as social media *check-ins*, shared *photos*, or (micro-)blog *posts*, on the other hand involve users actively choosing to contribute data pertaining to their social interactions in the physical world. These data have been analyzed in numerous ways with the goals of predicting human activity patterns (Scellato *et al*. 2011), identifying trends of temporal patterns for points of interest in cities (Sparks *et al*. in this issue), understanding human emotions from facial expressions (Svoray *et al*. 2018, Kang *et al*. 2019) or human sentiment at different neighborhoods from textual reviews (Hu *et al*. 2019), place recommendation systems (Xu *et al*. 2018), and urban visualization applications (McKenzie *et al*. 2015). While still considered part of the social sensing framework, these data are substantially different, and arguably more biased, than those contributed passively.

With the advancement of drive-by sensors, computer vision and deep learning techniques, street-level images become a new data source for understanding the physical environments and social environments. It enables the visual representation and exploration of urban environments using semantically segmented scene elements (Zhang *et al*. 2018). Spatiotemporal human activity information such as traffic flow (Zhang *et al*. 2019), neighborhoods demographic information (Gebru *et al*. 2017), and human perceived safety in cities (Li *et al*. 2015) can be inferred from street view images. In addition, street view images and 3D building models provide data support in urban design and planning. For example, a street-frontage-net (SFN) deep learning method has been developed to classify urban street-level images and evaluate the quality of street frontage from blank to active levels (Law *et al*. in this issue).

5. Datasets and reproducibility

Advancing GeoAI research requires high-quality geospatial datasets. Many AI models, particularly deep neural networks, need to be trained on a large set of well-labeled training data. It has long been recognized in the machine learning community that the quality of models follows the 'garbage in, garbage out' principle, i.e. a trained model is only as good as the quality of the training data. From this perspective, data are no longer merely resources to be mined by computational tools but are becoming part of the tools. High-quality datasets, such as ImageNet (Deng *et al*. 2009), have become critical enablers for the development of new AI methods. The domain of geography is fortunate to have many datasets of high quality in the public domain, such as the National Land Cover Dataset (NLCD) from the US Geological Survey and the American Community Survey (ACS) data from the US Census, not to mention the many available remote sensing images, global digital elevation models (DEM), and National Hydrography Datasets (NHD). With the change of data culture, an increasing number of companies are also sharing their geospatial data, such as the U.S. building footprint data by Microsoft, points of interest (POI) data by Yelp, and vehicle trajectory data by Uber and Didi. These and other shared geospatial datasets can become useful resources for developing future GeoAI models.

From a perspective of reproducibility and replicability, sharing the dataset based on which a GeoAI model was developed is necessary for other researchers to reproduce or replicate the model described in a research paper. According to Bollen *et al*. (2015), *reproducibility* refers to the ability of other researchers to duplicate the results of a prior study using the same data and procedures, while *replicability* refers to the ability of duplicating the results of a prior study using the same procedures but new data. As data are becoming integral to the models, one simply cannot arrive at the same model without having access to the original dataset used by the authors of said model. Sharing datasets on publicly accessible repositories, however, puts extra burdens on researchers, since a cleaned, well-organized, and carefully-documented repository of dataset requires significant additional effort that is often not rewarded in the current academic evaluations. Besides, there can be policies and privacy concerns that impede the effective sharing of datasets. Nevertheless, sharing a small sample of anonymized dataset can already go a long way toward enhancing the reproducibility and replicability of GeoAI research. The source code of the used architecture can be shared in a similar way, since the performance of deep neural networks is often affected by implementation details, such as random seeding and parameter initialization strategies. In this regard, we are glad to see that the papers in this GeoAI issue also shared a link to their GitHub repository with annotated data and source code.

Two directions could be explored to promote dataset and code sharing with the goal of supporting reproducibility and replicability in GeoAI research. First, we may continue enhancing our spatial data infrastructures (SDI) which serve as central platforms for sharing geospatial resources. Research efforts could be put on facilitating the search and discovery of resources on SDI (Hu *et al*. 2015), providing guidance on the best practices of data sharing, and designing automatic methods for improving the quality of geospatial data and metadata. Second, we could encourage the coupling of research articles and datasets in top journals of our domain. While this can be done by sharing a publicly accessible link of the repository within an article, an existing journal might offer a *dataset track*, or a new journal could be established specifically for publishing descriptions of geospatial datasets. There are already such dataset journals outside the domain of Geography, such as *Scientific Data* published by the Nature Publishing Group. These journal articles can give more credits to researchers who spent time and efforts to carefully collect, clean, and share datasets. On the other hand, new challenges need to be addressed on how to effectively review these dataset-description papers and how to ensure the quality and maintenance of the shared datasets.

6. Moonshots

Ideally, research around GeoAI and spatial data science more broadly would be focused around a few grand challenges. Such moonshots play an important role in measuring the progress of a community, explaining to others how some specific research direction contributes to a bigger picture, and agreeing on a common set of priorities. Here we outline one such moonshot.

Can we develop an artificial GIS analyst that passes a domain-specific Turing Test by 2030? Put differently, can we design a software agent that takes a user's GIS-related domain question, understands how to gather the required data, how to analyze them, and how to present the results in a suitable form? Imagine a user asking for available undeveloped spaces that are most suitable for community-based solar panel installations. The artificial GIS analyst would find the required data layers using SDI, perform operations such as insolation analysis, and return a suitability map, thereby (ideally) becoming indistinguishable from a human analyst. The key here is to open up GIS to Siri-like interaction for the masses, not to replace highly-trained GIS analysts performing complex analysis. Several large-scale projects such as the NSF funded EarthCube and, more recently, the Open Knowledge Networks track of NSF's Convergence Accelerator address the challenge of designing human and machine-readable and reasonable data repositories. Methods-wise, the above-mentioned research by Scheider *et al*. (2019) and other research teams can already be seen as contributions to this moonshot. From the industry's side, companies such as Esri have long experimented with automatically suggesting analysis and visualization options for common datasets and types. Hence, while ambitious, the envisioned artificial GIS analyst is a realistic goal if GeoAI research continues at today's speed.

7. Summary and conclusions

In this editorial, we motivated the need for GeoAI research and reviewed its origins. We have outlined three significant research directions, namely spatially explicit models, question answering, and social sensing, discussed the need for high-quality datasets and improved reproducibility, and presented a GeoAI moonshot as an example of a shared vision for the next ten years. We also hope that GeoAI and spatial data science more broadly will bring closer together the multitude of domains that work on or with spatiotemporal information. Finally, we believe that ethical consideration should be an essential part of responsible GeoAI research, both on the level of individual researchers as well as the community as a whole. We believe that the breadth of topics and techniques in this special issue (Acheson *et al*. in this issue, Zhu *et al*. in this issue, Sparks *et al*. in this issue, Law *et al*. in this issue, Li and Hsu in this issue, Duan *et al*. in this issue, Ren *et al*. in this issue, Guo and Feng in this issue, Xie *et al*. in this issue) is well representative of the current state-of-the-art in GeoAI.

Note

1. From personal communication in Spring 2017.

Acknowledgments

Authorship of this paper includes an employee of UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. Accordingly, the United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. Krzysztof Janowicz acknowledges support by the National Science Foundation under award 1936677 *'Convergence Accelerator Phase I (RAISE): Spatially-explicit Models, Methods, and Services for Open Knowledge Networks'*. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Grant McKenzie acknowledges support from the Natural Sciences and Engineering Research Council of Canada. Song Gao acknowledges support from the National

Science Foundation (Grant No. 1940091) and the Wisconsin Alumni Research Foundation. Finally, the authors would like to thank Gengchen Mai and Bo Yan for sharing comments and insights.

Funding

This work was supported by the National Science Foundation [1936677,1940091];Natural Sciences and Engineering Research Council of Canada;Wisconsin Alumni Research Foundation.

References

Acheson, E., Volpi, M., and Purves, R.S., in this issue. Machine learning for cross-gazetteer matching of natural features. *International Journal of Geographical Information Science*, 1–27.

Adams, B., *et al*., 2014. Geographic information observatories for supporting science. *In*: K. Janowicz, B. Adams, G. McKenzie, and T. Kauppinen, eds. *Proceedings of the workshop on geographic information observatories*, 23 September 2014 Vienna, Austria, 32–39.

Adams, B., McKenzie, G., and Gahegan, M., 2015. Frankenplace: interactive thematic mapping for ad hoc exploratory search. *In*: *Proceedings of the 24th international conference on world wide web*. International World Wide Web Conferences Steering Committee. Florence, Italy: ACM, 12–22.

Aggarwal, C.C. and Abdelzaher, T., 2013. Social sensing. *In*: A. C. Charu, ed. *Managing and mining sensor data*. Springer, 237–297.

Arora, S., Liang, Y., and Ma, T., 2017. A simple but tough-to-beat baseline for sentence embeddings. *In*: *Proceedings of the 5th international conference on learning representations*. Palais des Congrès Neptune, Toulon: ACM.

Ballatore, A., Bertolotto, M., and Wilson, D.C., 2013. Geographic knowledge extraction and semantic similarity in openstreetmap. *Knowledge and Information Systems*, 370 (1), 61–81. doi:10.1007/s10115-012-0571-0

Battaglia, P.W., *et al*.. 2018. Relational inductive biases, deep learning, and graph networks. *arXiv Preprint arXiv:1806.01261*.

Bollen, K., *et al*., 2015. Social, behavioral, and economic sciences perspectives on robust and reliable science. *Report of the Subcommittee on Replicability in Science Advisory Committee to the National Science Foundation Directorate for Social, Behavioral, and Economic Sciences*, 3, 4.

Chaix, B., 2018. Mobile sensing in environmental health and neighborhood research. *Annual Review of Public Health*, 39, 367–384. doi:10.1146/annurev-publhealth-040617-013731

Chu, G., *et al*., 2019. Geo-aware networks for fine grained recognition. *arXiv Preprint arXiv:1906.01737*.

Couclelis, H., 1986. Artificial intelligence in geography: conjectures on the shape of things to come. *The Professional Geographer*, 380 (1), 1–11. doi:10.1111/j.0033-0124.1986.00001.x

Deng, J., *et al*., 2009. Imagenet: a large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 248–255. doi:10.1037/a0016184

Duan, W., *et al*., in this issue. Automatic alignment of contemporary vector data and georeferenced historical maps using reinforcement learning. *International Journal of Geographical Information Science*, page in this issue.

Gao, S., *et al*., 2017. A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science*, 310 (6), 1245–1271.

Gebru, T., *et al*., 2017. Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, 1140 (50), 13108–13113. doi:10.1073/pnas.1700035114

Gil, Y., *et al*., 2007. Examining the challenges of scientific workflows. *Computer*, 400 (12), 24–32. doi:10.1109/MC.2007.421

Goodchild, M., 2001. Issues in spatially explicit modeling. *In*: D. C. Parker, T. Berger, and S. M. Manson, eds. *Agent-based models of land-use and land-cover change report and review of an international workshop october 47*. Irvine, CA, 12–15.

Goodchild, M.F. and Janelle, D.G., 2004. Thinking spatially in the social sciences. *Spatially Integrated Social Science*, 3–17.

Goodfellow, I., *et al*., 2014. Generative adversarial nets. *In*: Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence and K. Q. Weinberger, eds. *Advances in neural information processing systems*. Palais des Congrès de Montréal, Montréal CANADA: Neural Information Processing Systems (NIPS), 2672–2680.

Guo, Z. and Feng, -C.-C., in this issue. Using multi-scale and hierarchical deep convolutional features for 3D semantic classification of tls point clouds. *International Journal of Geographical Information Science*, 1–20.

Hey, A.J., *et al*., 2009. *The fourth paradigm: data-intensive scientific discovery*. Vol. 1. WA: Microsoft research Redmond.

Hu, Y., *et al*., 2015. Metadata topic harmonization and semantic search for linked-data-driven geoportals: a case study using arcgis online. *Transactions in GIS*, 190 (3), 398–416. doi:10.1111/tgis.12151

Hu, Y., 2018. Geo-text data and data-driven geospatial semantics. *Geography Compass*, 120 (11), e12404. doi:10.1111/gec3.12404

Hu, Y., Deng, C., and Zhou, Z., 2019. A semantic and sentiment analysis on online neighborhood reviews for understanding the perceptions of people toward their living environments. *Annals of the American Association of Geographers*, 1090 (4), 1052–1073. doi:10.1080/24694452.2018.1535886

Jacobs, N., *et al*., 2009. The global network of outdoor webcams: properties and applications. *In*: *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*. New York, NY: ACM, 111–120.

Janowicz, K., *et al*., 2014. Towards geographic information observatories. *In*: K. Janowicz, B. Adams, G. McKenzie, and T. Kauppinen, eds. *Proceedings of the workshop on geographic information observatories*, 23 September 2014 Vienna, Austria, 1–5.

Janowicz, K., *et al*., 2015, Mar. Why the data train needs semantic rails. *AI Magazine*, 360 (1), 5–14. doi:10.1609/aimag.v36i1.2560

Janowicz, K., *et al*., 2019. Using semantic signatures for social sensing in urban environments. *In*: *Mobility patterns, big data and transport analytics*. Amsterdam, Netherlands: Elsevier, 31–54.

Jones, J., *et al*., 2014. Making the web of data available via web feature services. *In*: *Connecting a digital Europe through location and place*. New York, NY: Springer, 341–361.

Kang, Y., *et al*., 2019. Extracting human emotions at different places based on facial expressions and spatial clustering analysis. *Transactions in GIS*, 230 (3), 450–480. doi:10.1111/tgis.12552

Khan, F., *et al*., 2019. Mobile crowdsensing: a survey on privacy-preservation, task management, assignment models, and incentives mechanisms. *Future Generation Computer Systems*, 100, 456–472. doi:10.1016/j.future.2019.02.014

Kuhn, W., 2012. Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information Science*, 260 (12), 2267–2276. doi:10.1080/13658816.2012.722637

Law, S., *et al*., in this issue. Street-frontage-net: urban image classification using deep convolutional neural networks. *International Journal of Geographical Information Science*, 1–27.

Li, M., *et al*., 2019. Reconstruction of human movement trajectories from large-scale low-frequency mobile phone data. *Computers, Environment and Urban Systems*, 77, 101346. doi:10.1016/j.compenvurbsys.2019.101346

Li, W. and Hsu, C.-Y., in this issue. Automated terrain feature identification from remote sensing imagery: a deep learning approach. *International Journal of Geographical Information Science*, 1–24.

Li, X., Zhang, C., and Li, W., 2015. Does the visibility of greenery increase perceived safety in urban areas? evidence from the place pulse 1.0 dataset. *ISPRS International Journal of Geo-Information*, 40 (3), 1166–1183. doi:10.3390/ijgi4031166

Liu, Y., *et al*., 2015. Social sensing: a new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 1050 (3), 512–530. doi:10.1080/00045608.2015.1018773

Mac Aodha, O., Cole, E., and Perona, P., 2019. Presence-only geographical priors for fine-grained image classification. *arXiv Preprint arXiv:1906.05272*.

Mai, G., *et al*., 2019a. Deeply integrating linked data with geographic information systems. *Transactions in GIS*, 230 (3), 579–600. doi:10.1111/tgis.12538

Mai, G., *et al*., 2019b. Relaxing unanswerable geographic questions using a spatially explicit knowledge graph embedding model. *In*: P. Kyriakidis, D. Hadjimitsis, D. Skarlatos, A. Mansourian, eds. *The annual international conference on geographic information science*. Berlin, Germany: Springer, 21–39.

Martin, H., *et al*., 2018. Graph convolutional neural networks for human activity purpose imputation. *In*: *NIPS spatiotemporal workshop at the 32nd Annual conference on neural information processing systems (NIPS 2018*).

McKenzie, G., *et al*., 2015. POI pulse: a multi-granular, semantic signature–based information observatory for the interactive visualization of big geosocial data. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 500 (2), 71–85. doi:10.3138/cart.50.2.2662

Moreau, L., *et al*., 2008. The provenance of electronic data. *Communications of the ACM*, 510 (4), 52. doi:10.1145/1330311.1330323

Openshaw, S. and Openshaw, C., 1997. *Artificial intelligence in geography*. 1st ed. New York, NY: John Wiley & Sons, Inc. ISBN 0471969915.

Polson, N.G. and Sokolov, V.O., 2017. Deep learning for short-term traffic flow prediction. *Transportation Research Part C: Emerging Technologies*, 79, 1–17. doi:10.1016/j.trc.2017.02.024

Regalia, B., *et al*., 2016. Crowdsensing smart ambient environments and services. *Transactions in GIS*, 200 (3), 0382–398. doi:10.1111/tgis.12233

Regalia, B., Janowicz, K., and McKenzie, G., 2019. Computing and querying strict, approximate, and metrically refined topological relations in linked geographic data. *Transactions in GIS*, 230 (3), 601–619. doi:10.1111/tgis.12548

Ren, Y., *et al*., in this issue. A hybrid integrated deep learning model for the prediction of citywide spatio-temporal flow volumes. *International Journal of Geographical Information Science*, 1–22.

Resch, B., *et al*., 2015. Urban emotionsgeo-semantic emotion extraction from technical sensors, human sensors and crowdsourced data. *In*: *Progress in location-based services 2014*. Berlin, Germany: Springer, 199–212.

Rodriguez, A.M. and Egenhofer, M.J., 2004. Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. *International Journal of Geographical Information Science*, 180 (3), 229–256. doi:10.1080/13658810310001629592

Scellato, S., Noulas, A., and Mascolo, C., 2011. Exploiting place features in link prediction on location-based social networks. *In*: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. San Diego, CA: ACM, 1046–1054.

Scheider, S., Ballatore, A., and Lemmens, R., 2019. Finding and sharing gis methods based on the questions they answer. *International Journal of Digital Earth*, 120 (5), 594–613. doi:10.1080/17538947.2018.1470688

Seif, H.G. and Hu, X., 2016. Autonomous driving in the icityhd maps as a key challenge of the automotive industry. *Engineering*, 20 (2), 159–162. doi:10.1016/J.ENG.2016.02.010

Silva, M.J., *et al*., 2006. Adding geographic scopes to web resources. *Computers, Environment and Urban Systems*, 300 (4), 378–399. doi:10.1016/j.compenvurbsys.2005.08.003

Smith, T.R., 1984. Artificial intelligence and its applicability to geographical problem solving. *The Professional Geographer*, 360 (2), 147–158. doi:10.1111/j.0033-0124.1984.00147.x

Sparks, K., *et al*., in this issue. A global analysis of cities geosocial temporal signatures for points of interest hours of operation. *International Journal of Geographical Information Science*, 1–18.

Svoray, T., *et al*., 2018. Demonstrating the effect of exposure to nature on happy facial expressions via flickr data: advantages of non-intrusive social network data analyses and geoinformatics methodologies. *Journal of Environmental Psychology*, 58, 93–100. doi:10.1016/j.jenvp.2018.07.006

Vahedi, B., Kuhn, W., and Ballatore, A., 2016. Question-based spatial computinga case study. *In*: *Geospatial data in a changing world*. Helsinki, Finland: Springer, 37–50.

Wang, M., *et al*., 2018. Towards empty answers in sparql: approximating querying with rdf embedding. *In*: *International semantic web conference*. Springer, 513–529. doi:10.1177/1753193417739248

Xie, Y., *et al*., in this issue. A locally-constrained yolo framework for detecting small and densely-distributed building footprints. *International Journal of Geographical Information Science*, 1–25.

Xu, L., *et al*., 2018. Check in or not? a stochastic game for privacy preserving in point-of-interest recommendation system. *IEEE Internet of Things Journal*, 50 (5), 4178–4190. doi:10.1109/JIOT.2018.2847302

Yan, B., *et al*., 2017. From itdl to place2vec: reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. *In*: *Proceedings of the 25th ACM SIGSPATIAL international conference on advances in geographic information systems*. Redondo Beach, CA: ACM, 35.

Yan, B., *et al*., 2018. xnet+ sc: classifying places based on images by incorporating spatial contexts. *In*: S. Winter, A. Griffin and M. Sester, eds. *10th international conference on geographic information science (GIScience 2018)*. Melbourne, Australia: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Yan, B., *et al*., 2019. A spatially explicit reinforcement learning model for geographic knowledge graph summarization. *Transactions in GIS*, 230 (3), 620–640. doi:10.1111/tgis.12547

Zang, A., *et al*., 2017. Accurate vehicle self-localization in high definition map dataset. *In*: S. Gao, J. Dai, G. He, Y. Xu, and X. Chen, eds. *Proceedings of the 1st ACM SIGSPATIAL workshop on high-precision maps and intelligent applications for autonomous vehicles*. Redondo Beach, CA: ACM, 2.

Zeile, P. and Resch, B., 2018. Combining biosensing technology and virtual environments for improved urban planning. *GI_Forum*, 1, 344–357. doi:10.1553/giscience

Zhang, F., *et al*., 2018. Representing place locales using scene elements. *Computers, Environment and Urban Systems*, 71, 153–164. doi:10.1016/j.compenvurbsys.2018.05.005

Zhang, F., *et al*., 2019. Social sensing from street-level imagery: a case study in learning spatio-temporal urban mobility patterns. *ISPRS Journal of Photogrammetry and Remote Sensing*, 153, 48–58. doi:10.1016/j.isprsjprs.2019.04.017

Zheng, Y., *et al*., 2014. Diagnosing new york city's noises with ubiquitous data. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. Seattle, WA: ACM, 715–725.

Zhu, D., *et al*., in this issue. Spatial interpolation using conditional generative adversarial neural networks. *International Journal of Geographical Information Science*, 1–24.

Krzysztof Janowicz
STKO Lab, Department of Geography, University of California, Santa Barbara, CA, USA
jano@geog.ucsb.edu

Song Gao
GeoDS Lab, Department of Geography, University of Wisconsin, Madison, WI, USA
http://orcid.org/0000-0003-4359-6302

Grant McKenzie
Platial Analysis Lab, Department of Geography, McGill University, Montreal, Canada
http://orcid.org/0000-0003-3247-2777

Yingjie Hu
GeoAI Lab, Department of Geography, University at Buffalo, Buffalo, NY, USA
http://orcid.org/0000-0002-5515-4125

Budhendra Bhaduri
Oak Ridge National Lab, Oak Ridge, TN, USA
http://orcid.org/0000-0003-1555-1377