Accelerating ethics, empathy, and equity in geographic information science

T. A. Nelson, M. F. Goodchild, and D. J. Wright

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2021. Contributed by D. J. Wright; received November 10, 2021; accepted February 22, 2022; reviewed by Sarah Elwood, Diana Liverman, and Nadine Schuurmann

Science has traditionally been driven by curiosity and followed one goal: the pursuit of truth and the advancement of knowledge. Recently, ethics, empathy, and equity, which we term “the 3Es,” are emerging as new drivers of research and disrupting established practices. Drawing on our own field of GIScience (geographic information science), our goal is to use the geographic approach to accelerate the response to the 3Es by identifying priority issues and research needs that, if addressed, will advance ethical, empathetic, and equitable GIScience. We also aim to stimulate similar responses in other disciplines. Organized around the 3Es we discuss ethical issues arising from locational privacy and cartographic integrity, how our ability to build knowledge that will lead to empathy can be curbed by data that lack representativeness and by inadvertent inferential error, and how GIScientists can lead toward equity by supporting social justice efforts and democratizing access to spatial science and its tools. We conclude with a call to action and invite all scientists to join in a fundamentally different science that responds to the 3Es and mobilizes for change by engaging in humility, broadening measures of excellence and success, diversifying our networks, and creating pathways to inclusive education. Science united around the 3Es is the right response to this unique moment where society and the planet are facing a vast array of challenges that require knowledge, truth, and action.

Significance

There has been a groundswell in the support needed to center ethics, empathy, and equity in scientific thought and practice. Drawing on our experience from GIScience, our goal is to accelerate ethical, empathetic, and equitable scientific practices. Many of the opportunities and challenges we outline are broadly applicable and will stimulate the conversations needed to accelerate transformation of science practice and culture. With an emphasis on practical suggestions for reshaping science, we invite all scientists to join in a fundamentally different approach. This paper is a step toward mobilizing the scientific community toward ethics, empathy, and equity by inviting humility, broader measures of excellence and success, diversity in our networks, and the creation of pathways to inclusive education.

Most scientific research has traditionally been fueled by the innate curiosity of the researcher, by the desire to generate replicable and generalizable knowledge, and by the need to address practical problems. While there are, of course, less-noble drivers of science, most often researchers nobly aim their tools at knowledge creation and measure contributions through publications, awards, the support of funding agencies and our institutions, and various less-tangible forms of personal satisfaction. Yet, today we find a groundswell of support for complementary norms which have been percolating in science for some time. Here we focus on the “three Es”—ethics, empathy, and equity—which have scientists reflecting on and responding to questions such as the following:

- Given that we strive at all times for scientific practices to be ethical, how can science foster empathy and equity as well?
- Will our syllabi and courses attract and address the needs of all of our students?
- Does the culture of science that has evolved over the centuries reflect the needs and desires of all of humanity, or is science better at serving some segments of society than others?

These fundamental questions and many more like them are already the focus of long-standing initiatives such as the NSF’s “Broader Impacts” criterion for research proposals. There are also professional development programs and networks seeking to accelerate progress in these areas, such as the GeoEthics project of the American Association of Geographers (https://aag-geoethics-series.secure-platform.com/a); the American Geographical Society’s EthicalGEO initiative (https://ethicalgeo.org/); the American Geophysical Union’s LANDiNG (Leadership Academy and Network for Diversity and Inclusion in the Geosciences, https://www.agu.org/AGU-LANDInG); the University Consortium for Geographic Information Sciences’ TRELLIS-GS (Training and Retaining Leaders in STEM-Geospatial Sciences, https://www.ucgis.org/trelis); the American Association for the Advancement of Sciences’ Project on Science, Technology, and Disability (https://www.aas.org/programs/education-and-human-resources/project-science-technology-and-disability); and NorthStar, for increasing the representation, belonging, and inclusion of people of African descent in the geospatial industry and academia (https://gisnorthstar.org/). There are inclusive professional organizations such as the Society for the Advancement of Chicanos/Hispanics and Native Americans in Science
systems for managing and tracking spatially distributed activities such as transit use or crime, for situation awareness in the military and intelligence communities, for assessing the effectiveness of public-health systems, for emergency management, and for modeling social and environmental systems (e.g., refs. 5 and 6). GIS is today an important research tool in virtually all of the social and environmental sciences (e.g., ref. 7), an essential administrative tool, and a key design tool in architecture, landscape architecture, and urban planning (8–10).

In 1992 Goodchild coined the term “GIScience” to encompass the rigorous scientific treatment of GIS design and methods (11); alternative terms with similar meaning include “geomatics,” “geoinformatics,” and “spatial data science.” GIScience is built on the idea that humans have specific ways of orienting themselves, acquiring and communicating spatial knowledge, and finding their way (12), and principles of spatial cognition are often important in the design of GIS. Elegant data structures and indexing schemes have been devised for spatial data (e.g., refs. 13 and 14), and powerful algorithms have been created for performing basic operations (e.g., refs. 15–18). Another unique aspect of GIScience is that it tends to adhere to specific principles, most notably spatial dependence and spatial heterogeneity (19), and is impacted by unique uses such as the modifiable areal unit problem (20). These challenges require specialized treatment of statistical inference and uncertainties (21).

Some of the first discussions of 3E issues in GIScience occurred in the late 1980s and early 1990s (e.g., ref. 22). GIS at that time was expensive and tended to be available only to governments and the military. At the same time, because of the limited levels of detail of many data sources and the limited power of computational systems, some GIS solutions could be seen as simplistic and naïve. GIS also allowed individual locations to be easily shared, raising questions about locational privacy (23). Pickles (24) and others argued that maps, at the time the primary source of much GIS data, are often to be interpreted as social constructions that represent the political, commercial, and other agendas of their makers, rather than as scientifically objective representations of reality. Social critiques of GIS, that evaluate GIS through the lens of critical theory, have highlighted the assumptions of the practitioners and shown how they are reflected in knowledge production (25, 26). These arguments had a significant impact on the development of GIS and on GIScience.

Presently, however, it is hard to imagine any conversation in GIScience, or indeed in any other scientific community, that does not consider some mix of ethics, empathy, and equity. The new focus on the 3Es is, in part, a signal of the moment. The Black Lives Matter movement has brought systemic racism to the forefront of many conversations. The broader field of geosciences, in which GIScience often sits, acknowledges collective underperformance in training and amplifying scholars and professionals in Black and other underrepresented minorities (27, 28). As well, in GIScience we are reexamining GIS as a tool to support more equitable and inclusive design and decision-making (29) and highlighting the dangerous applications of GIS technology such as using spatial databases for policing (30). COVID-19 has also contributed to our focus on the 3Es, as GIS has been critical to the COVID-19 response (31). Discussions about how to build and implement spatial technology have highlighted the need to balance privacy and analytical power when working with locational data. For example, during the early days of COVID-19, spatial technology for digital contact tracing was an active topic of discussion. However, as the pandemic continued, we were reminded that widespread use of any technology requiring knowledge of individual
locations is difficult to implement, and we now recognize the hazards associated with positional uncertainty.

Our primary goal in this paper is to accelerate response to the 3Es in science by identifying priority issues and research needs that, if addressed, will advance ethical, empathetic, and equitable GIScience and to stimulate similar responses in other disciplines. Action in GIScience will have broad implications for all science as spatial data, methods, and software have permeated across science, including to engineering (32), health science (33), ecology (34), and social sciences (35). We recognize that issues and potential responses to the 3Es are a vast canvas. In GIScience a sample of the relevant topics includes ethics (36–38); data representation (39); justice, equity, diversity, and inclusion (40–44); location and privacy (45); inference from spatial data (46); provenance and uncertainty in data (47); teaching practices (48, 49); and reproducibility and replicability (50–54). As we cannot give adequate attention to all topics, we highlight those that either require a unique response from GIScience or where leadership from GIScience will have broad impacts for science, academia, and society. Topics are organized around each of the 3Es (Fig. 1).

**Ethics of Location**

As pertaining to science, ethics typically involves reflection upon moral questions that arise in research, publication, data collection and analysis, and other professional activities, all governed by a set of moral principles or codes of behavior (55). While ethics in science includes everything from human subjects to cultures of inclusion, ethics intersects uniquely with the field of GIScience due to the nature of data, the power of cartographic visualization, and the impact of those topics that attract the gaze of geographers. Professionalism in GIScience (pertaining to GIS, but also remote sensing with satellites, drones, and smaller IoT sensors, cartography, and quantitative spatial analysis) now involves a commitment to ethical practice as informed by a more sophisticated understanding of the ethical implications of GIS technologies. For example, the use of GIS for military and surveillance purposes, the lack of privacy introduced by mobile mapping devices, and the use or misuse of GIS for conservation and sustainability continue to be challenging issues and topics of deep concern for many. There are examples of rich studies of ethics in GIScience (mentioned prior), yet, as a broader academic subfield, ethics has been on the periphery and often treated as a footnote, as the final lecture in an otherwise technical course (49), or as the sole course within an entire graduate degree or certificate program [https://www.e-education.psu.edu/research/projects/gisethics (48)].

On the GIS industry front, however, there are established and published codes of ethics and rules of conduct (two different things) that must be agreed to as a final step in professional certification by organizations such as the GIS Certification Institute and the American Society for Photogrammetry and Remote Sensing. In addition, the Urban and Regional Information Systems Association was one of the first scholarly and industry-focused nonprofits to publish a code of ethics (https://www.urisa.org/about-us/gis-code-of-ethics/), and the American Association of Geographers, which includes scholarly specialty groups in cartography, cyberinfrastructure, digital geographies, geographic information science and systems, and remote sensing, has frequently revised its Statement on Professional Ethics up to the time of this writing (https://www.aag.org/cs/about_aag/governance/statement_of_professional_ethics).

The success of GIScience in the 21st century has brought an even greater urgency and complexity to ethical conversations. GIScience is has infused many arenas, and as a result ethics in GIScience has far reaching impacts. Everything from our academic activities to the individual privacy of everyday smartphone users are impacted by the technology and the norms of our field. Ethics is finally encoded in the latest version of the University Consortium of Geographic Information Science’s Geographic Information Science & Technology Body of Knowledge (56, 57). While institutionalized ethics is an important starting point, codes will be implemented by humans and, therefore, must be accompanied by diverse teams. Two ethical concerns on the forefront of our minds relate to the abundance and growing use of locational data and maps: locational privacy and cartographic integrity.

**Locational Privacy.** Humans have traditionally used place names to refer to location: names of rivers, lakes, mountains, and other natural features; names of cities, regions, streets, and neighborhoods; and names of businesses and institutions. The accurate measurement of location was an expensive process requiring the use of sophisticated instruments and technical skill. By 2010, however, low-cost GPS chips marketed as standard smart-phone components made it possible to determine location almost instantly and to a few meters and led to the development of numerous location-based services: search for nearby businesses, wayfinding, ride hailing, and the operation of autonomous vehicles. In some cases, it was also possible for the providers of such services to acquire the locations of their users and to aggregate and market such information to third parties (58). In principle these practices occurred only when users opted in; in practice, however, many users paid little attention to the detailed terms and conditions of these services.

Dobson and coauthors (59, 60) coined the term “geoslavery” to describe what they saw as the social consequence of this kind of surveillance, “a practice in which one entity, the master, coercively or surreptitiously monitors and exerts control over the physical location of another individual. Indeed, there are many examples of automated surveillance being used for racial profiling [e.g., refs. 61 and 62]]. Enhanced surveillance and control may be attained through complementary monitoring of functional indicators such as body temperature, heart rate, and perspiration.” Complementary monitoring can also be provided by the phone’s accelerometer and its route-following apps, and some models of new cars include direct internet connections that allow the car’s speed and acceleration to be remotely monitored, whether or not a smart phone is present. It is not hard to envision the consequences of this kind of capability. Today, for example, State Farm Insurance offers two programs that

---

**Fig. 1.** Organization of the paper around the 3Es: ethics, empathy, and equity.
reward drivers with reduced insurance rates based on the company’s monitoring: Steer Clear, which “reinforces good driving habits in young drivers and helps them earn a valuable discount” and Drive Safe & Save, which “collects basic information about your driving that may earn you a discount and save you some money” (https://www.statefarm.com/customer-care/download-mobile-apps/state-farm-mobile-app).

Location data from smart phones have frequently been identified as the basis for COVID-19 surveillance. It is possible, in principle, to detect when two smart phones come within a given distance of each other for a given period of time (say, 2 m and 15 min, respectively, enough for COVID infection to pass between the owners of the phones). GPS is an obvious way to do this, though uncertainty in positioning will lead to numerous false positives and false negatives. Early in the COVID outbreak Apple and Google announced a joint program to use Bluetooth technology for contact tracing, with advantages in terms of personal privacy (63), but while this may appear to be an elegant technical solution to contact tracing, the recent “pingdemic” experience in the United Kingdom points to some of its less-attractive consequences (64).

Some programs of statistical data collection have dealt with the problem of locational privacy by aggregating data to reporting zones such as counties or census tracts, blocking any access to individual data. Randomly distorting locations, or geomasking, is another possible approach. Individual data can be anonymized, but it is often trivially easy to deanonymize data by linking together the “pings” from a single device to reconstruct a track (e.g., refs. 58 and 64). Privacy of health data, including the United States’ Health Information Portability and Accountability Act (HIPAA) of 1996 (https://www.hhs.gov/hipaa/index.html), limits use on GI in the health sector, and some jurisdictions have enacted other limits. However, if the data and methods of GIS-science are to be used ethically, we clearly need to advance norms and approaches to protecting locational privacy in all sectors. Collection of individual-level location data has outpaced development guidelines for ethical use, and a broad spectrum of responses to locational privacy has emerged. On the one hand, research institutions require IRB (Institutional Review Board) reviews that place strict limitations on the use of fine-resolution locational data, while on the other hand new commercial industries have developed around the largely unregulated collection and sale of individual location data. The popularity of GI, combined with inconsistency in norms regarding ethical use, has led to new initiatives like the American Geographical Society’s EthicalGEO initiative (https://ethicalgeo.org/) and the American Association of Geographers’ GeoEthics program, both launched in 2019 (65), which are leading an international dialogue on standards of use for locational data.

**Cartographic Integrity.** With the proliferation of digital interactive mapping and analysis platforms, many of them cloud-based, there is also a need to update the ethical guardrails for cartography, a foundation of GIScience. We have seen an explosive increase in the use of maps in society, no longer solely in the domain of professional cartographers. With every public-health agency and major news media outlet serving COVID-19 data through both static and interactive maps and dashboards, we are reminded of the cartographer’s role in shaping public opinion and informing policy. The growing reach and influence of maps raises age-old questions (e.g., ref. 66) of how cartography can be used as a mechanism to generate knowledge or be manipulated to tell a particular story. For example, colleges and universities have over the past year released a plethora of coronavirus dashboards, many of them with interactive maps, and listing active cases, positivity rates, test numbers, and other statistics (Fig. 2). However, as Vasquez (67) points out, without a uniform standard for what the maps and other information presented on these dashboards should include or how often to update them, confusion reigns as to how truly bad a campus coronavirus outbreak may be—and where—on any given day. Institutions are searching for an appropriate level of transparency, given concerns over student privacy, as well as a desire to protect their reputations and their ability to retain or attract students, faculty, and donors, all with the lives of the people already within their communities in the balance. On the industry front, an emerging question is, Does your workforce truly reflect your community? Accordingly, workforce equity dashboards are coming to the fore accordingly to help answer this question. (Fig. 3)

Monmonier’s (68) instant classic of cartographic literacy and integrity, *How to Lie with Maps*, as well as a related paper (22), set an early standard, especially as a caution to map users as to how mapmakers can either accidentally or intentionally distort the information that is presented on a map. As we have moved further into the digital age, including the advent of virtual and augmented reality, and the machine- and deep-learning inherent in AI, we have become aware of the myriad possible distortions not only of the underlying data but of statistics and other analyses performed on the data and their final depiction in a digital cartographic product. The third edition of his book (69) includes new cautionary tales regarding online mapping, with the accompanying opportunities for “cartographic mischief,” deceptions by way of AI and physics-based algorithms that can add degrees of fake “reality” to maps and satellite images (70), location spoofing (interference with legitimate GPS signals or falsifying the true location of a device), and even map propaganda.

Field (71), in *Thematic Mapping: 101 Inspiring Ways to Visualise Empirical Data*, explores the issue of cartographic integrity for the 21st century, parallelizing the discussions of Monmonier (69) and citing in particular the human factors of oversight, poor judgment, lack of cartographic training, and, perhaps most importantly, a persistent reliance on what may be inappropriate software defaults. These failings often occur even if the mapmaker has the noblest of intentions. Others have also written on the topic of ethics in cartography (e.g., refs. 72–75), but Field (71, 76) chose to update the guidelines of Dent et al. (77) with accessible principles such as have a straightforward agenda and purpose; strive to know your audience (the map reader); do not intentionally lie with data; show all relevant data whenever possible; do not discard data because it might be contrary; strive for accurate portrayal of the data; report all data sources and errors; ensure that symbols do not bias the interpretation of the map; ensure that the map is reproducible by others; be attentive to differing cultural values and principles; and, once again, do not let software defaults drive your design.

This tome on contemporary cartographic practice also stands apart in that the emphasis on “manners for the modern mapmaker,” whether professional or not, is coupled with an annual and free massive open online course, or MOOC, based directly on the content of the 2018 book but also continually updated (78). Enrollment in this MOOC since its 2018 inception has numbered in the tens of thousands globally (completion rates do vary, as with most MOOCs). It is resources such as these that indicate the continuing desire of cartographers to strive for truth, fairness, and accuracy in their mapmaking efforts, including those online. Such resources seek to attract and address the needs of a broader cadre of students and more segments of society. They are sorely needed in a posttruth world, filled with...
alternative facts, fake news, and conspiracy theories that erode trust in science, hampering healthy curiosity in both research and practice, and ultimately stymying our ability to produce maps that can be fully trusted, especially when lives are at stake.

From Ethics to Empathy

If ethics, in part, is a reflection upon moral questions that govern codes of human behavior in professional scientific practice, empathy has long been recognized as the driving force behind much human behavior. Like many science fields, GIScience has been a vehicle for building empathy through generating new understanding of people and place. However, among science approaches, GIScience is perhaps uniquely placed to support building empathy, by enabling integration of data and visualization of multiple perspectives. There are many examples, for instance, of humanitarian groups leveraging the power of GIS to raise awareness and support logistics (e.g., ref. 79). As well, participatory GIS, which enables community engagement in mapping and analysis, creates important mechanisms for empowering underserved and diverse groups (80) and applications of Storymaps, an easy-to-use tool for integrating maps and visualize with narratives, is putting powerful communication tools in the hands of many (81, 82).

While empathy has long been an outcome of knowledge, only recently has empathy been identified as a driving force for research and learning, stemming from one’s curiosity or desire to learn about the experiences of others (83). Traditionally, curiosity has been the driver of research and empathy an outcome that results from creation of knowledge and enables us to take on new perspectives (84). Recently, we have observed a shift toward empathy as more of a driver of research and as the impetus for curiosity, at the same time enhancing skills in critical thinking, creative thinking, and collaborative discovery (83). Empathy-led curiosity is bringing meaning to the academic work of many students and scholars and is both an individual and collective response to systemic racism, social disparities, climate-change impacts, and the global pandemic.

Simultaneous to the increase in empathy as a driver of research, there has been growing abundance of diverse spatial data and methods. On the surface, more data and methods seem inherently desirable and could help to advance and diversify knowledge. Yet, we see a potential tension between the goal of advancing empathy through research that is inclusive, on the one hand, and the nature of some emerging data and approaches on the other. Specifically, we consider how two aspects of GIS may be inadvertently interfering with our aims to improve empathy: data representativeness and inference.

Data Representativeness. Abundant spatial data are creating new opportunities for analysis and decision-making support. Unlike data collected systematically using predetermined sampling frameworks, much of the increase in spatial data availability is associated with opportunistic data collection from GPS-enabled cell phones, social media, and programs that make use of volunteered geographic information (VGI). For example, point-of-interest datasets, like SafeGraph, compile data on movement from millions of cell phones (85); Strava-generated datasets from application users track bicycling and other athletic activities (39); and OpenStreetMap (OSM) is arguably the most successful VGI project, enabling crowdsourced mapping globally (86) and often more complete and more timely than official data sources (e.g., ref. 87).

There are many benefits to big spatial data and opportunistic data collection, but the trade-off is sampling bias. New sources of spatial data are creating representations of phenomena that were previously unmapped, with spatial and temporal resolutions that are incredibly fine. Yet, sampling is biased toward people who have access to technology or interest in contributing data. In the example of SafeGraph, people without phones will be missing from the sample, and this most likely includes children, older adults, people who are homeless or living in poverty, and people with disabilities. In the cases of StravaMetro and OSM, men contribute the majority of data (88, 89). Use of VGI and other opportunistically generated geographic data raises...
concerns about the implications of using data that undersample people already experiencing barriers and underservice. If data only represent the dominant culture, they cannot be used to generate knowledge or empathy for people experiencing barriers to gaining economic, social, and political capital.

Data integration and modeling can be approaches to reducing bias in crowdsourced and VGI datasets. GIS data integration uses location as a universal indexing system and allows multiple and diverse datasets to be combined to create more comprehensive and inclusive datasets. As a demonstration, consider an example from transportation safety, where official crash reports, representing <20% of bicycle crashes (90), can be combined with VGI [e.g., https://bikemaps.org/(91)] on crashes and near misses to provide a more complete map of road safety (92). In this case, both official and VGI data sources are missing information but in combination are more representative. As well, modeling can be used to reduce bias in opportunistically sampled GIS data. Using training or truth data, it is possible to build statistical relationships with a sample of data that may be biased and thus predict spatial variation error or bias (e.g., ref. 93). Use of training data to classify a sample has been widely applied in remote sensing for image classification; rather than build labels from reflectance values, training data can be used to quantify spatial variation in bias.

While data integration and modeling can lead to more representative spatial data on individual projects, to collectively advance the 3Es the field of GIScience needs a framework for identifying, estimating, and communicating bias in spatial data. Researchers and funding agencies could prioritize research that will build a generalizable framework for understanding sampling bias and representativeness in spatial data. Such a framework might help us to move toward consistent ways of documenting and reporting sampling bias, to develop measures for evaluating change in data bias and representativeness, and to build a common language for dialogue on how sampling bias limits the generalizability of our findings. Evaluating assumptions and bias has always been a critical part of scientific peer review. The shift toward developing scientific evidence from data that are generated for nonscience purposes should be a focus of reviews, and having a framework and language for understanding the structures of emerging datasets will allow us to build toward theory. Data scientists are considering error frameworks for nonspatial data, and while these can provide a roadmap, additional dimensions and characteristics of map data will require more nuanced frameworks. Representativeness in data will likely lack a universal solution; rather, there will patterns reflective of different social and geography contexts from which data emerge. GI solutions may be best focused on metalevel frameworks for understanding the link between patterns of representation and the contextual processes that drive them.

Inference. Much empirical work in GIScience, as in all science, relies on the making of inferences from conditions as they are observed on the ground. These so-called natural experiments differ in profound ways from the controlled experiments of the laboratory. For example, the principle of spatial heterogeneity implies that inferences made from observations in one geographic area will not necessarily replicate inferences from other geographic areas (53). When inferences are drawn from spatially aggregated data, such as data about counties or census tracts,
the modifiable areal unit problem (94) asserts that those inferences will change when the spatial units change—results from an analysis of counties will not match results from an analysis of the same data aggregated to voting districts, for example.

Many of the most compelling results of spatial analysis are based on inferences drawn from such cross-sectional data, that is, geographic data obtained at one point in time. Unfortunately, the principle of equifinality asserts that the same cross-sectional form can result from multiple distinct processes; thus, analysis of the same data can lead to many alternative inferences, some perhaps more empathetic than others. In research practice, then, it may be tempting to claim that the data support a specific inference that favors some preconceived hypothesis, belief, or argument, while other possible inferences are ignored. If the goal of inference is the generation of knowledge, then inferential errors are a barrier to the kinds of understanding needed to grow empathy. The link between inference and knowledge is not new, nor is it specific to GIScience, yet as spatial data have proliferated and the use of black-box approaches to analysis have gained traction we see greater temptation and opportunity for inadvertent inferential error and damage to the goal of empathy.

An excellent illustration of the need for inferential caution is the well-known spatial analysis example of Dr. John Snow and the cholera outbreak in London in 1854 (95). At the time the popular belief was that cholera was transmitted through miasma, that is, polluted and noxious air. Snow’s map (Fig. 4) clearly showed a spatial association between the locations of the more than 700 deaths in the outbreak and a water pump located in Broad Street. This natural experiment led Snow to the inference that cholera was passed through drinking polluted water, not miasma, and a controlled experiment (preventing access to the water from the pump) provided confirmation. Note, however, that based solely on the map one could infer that the miasma hypothesis was correct (the bad air being limited to the densely populated and socially deprived Soho neighborhood), and the map could also support a contagion hypothesis (the original carrier being located near the pump), or even a demographic hypothesis (the pattern of deaths reflects the presence of some population at risk—perhaps the elderly). The three alternative inferences might be regarded as more complex, so the principle of Occam’s razor might be cited as additional support for Snow. Today, Snow’s control over access to the pump would likely not be possible. Without the ability to conduct an intervention, even the geographer’s favorite example of spatial pattern and inference could have been problematic. It is a reminder that we must be cautious in our efforts at inference and be mindful that massive amounts of data will not necessarily overcome this challenge.

While all disciplines are challenged to design studies carefully to support inference, spatial data create some unique concerns, of which the ecological fallacy provides an example. Suppose a positive correlation is observed between some demographic factor, such as race, and the incidence of infection from some organism, using aggregate data at the level, say, of the census tract. It is tempting but nevertheless fallacious in such a situation to infer that race is a risk factor in infection. Instead, the association has merely demonstrated that high rates of infection tend to co-occur with particular racial patterns. King (96) has argued that inferences can be made at the individual level from aggregate data, but only if certain mathematical assumptions can be made.

The techniques of statistical inference have their origins in the work of agronomists and experimental psychologists, where research is often conducted on samples of individuals or plants that have been drawn randomly and independently from larger populations. These techniques then allow inferences to be made about the larger population—in short, to generalize away from the sample. When these methods are applied to the natural experiments of GIScience two problems arise. First, the principle of spatial dependence makes it very likely that the individuals in a sample will be correlated, or more technically that the number of degrees of freedom has been inflated. Many methods of geostatistics and spatial statistics have been developed to deal with this issue. Second, it is common in GIScience for research to deal not with a randomly chosen sample but with the entire available set: all of the counties in the United States, for example, or all of the census tracts in a city. In such situations it is difficult to conceive of a larger population, difficult to see the set as randomly chosen from it, and difficult therefore to generalize away from the study.

In summary, inference in GIScience is often, if not always, problematic, and the investigator will frequently be faced with a number of alternative inferences. In some cases, the choice will be driven by some prior hypothesis or belief, and in other cases by the use of Occam’s razor. These practices, however, are likely to be disrupted in various ways by the 3Es. Given the power of spatial analysis to inform, and the ambiguities that often arise when inferences are drawn, our field requires ethical guidelines to provide checks and balances on how we use data, methods, and inference. Without ethical guardrails we run the risk of amplifying disparities through false findings and misrepresentation. These risks are ubiquitous to science, but they are especially worrisome in spatial science where the additional dimensions of data (x, y, and t) create ample opportunities for uncertainty.

**From Ethics and Empathy to Equity**

With professional scientific practice governed by codes of desirable behavior, driven further by a sincere, empathy-led curiosity to bring societal meaning to academic work, it follows that a major outcome of that work would be to achieve equity. Equity
can be said to exist when everyone is able to achieve equal outcomes through equal access to opportunities and resources (97) and to foster a clear sense of belonging once those opportunities or resources are obtained. It is further defined by the Annie Casey Foundation (98) as an inability to predict outcomes based solely on race, ethnicity, or other defining characteristics, as for example when no group has a clear advantage in predicted high-school graduation rates. Racial equity in particular means that a person's opportunities are not determined by their physical appearance or racial background. As such, equity differs from equality. While equality exists when everyone is treated in the same way, equity places the emphasis on potential, and occurs when everyone is able to achieve the same outcomes. Equity always considers past history and acknowledges the plethora of existing systemic barriers, with the aim of lifting everyone to the same level, again with that sense of belonging. The Black Lives Matter movement has created a new urgency and more willingness to respond to calls for equity in science, academia, and society (e.g., refs. 43 and 99–101). Authentically supporting equity requires change. The issue is complex and will be a challenge, but it is high time. In this section we discuss the potential role of GIScience in achieving equity through social justice and its potential for democratizing access to information technology.

Social Justice. As an interdisciplinary field covering physical and social sciences as well as digital humanities, GIScience is increasing its emphasis on equity and social justice, perhaps more than most other technology-centric disciplines. Social justice can be defined as the mechanisms by which we construct equality (102) and is closely related to equity, in part, through the connection between race and space as key predictors of inequity (e.g., ref. 103). It follows that the reliance on GI and associated methods and tools to represent and analyze issues of social justice such as environmental racism, homelessness, income inequality, unemployment and housing inequality (especially due to COVID), opioid abuse, blight, community policing, access to public transit, food deserts and more, including the derivation of actual quantitative indicators, is necessary, logical, and appropriate (97).

Achieving social justice requires in part an understanding of how current systems have created inequities and injustices over time, and once understood we can identify the role we play within the system and what we can do to change it. By collecting knowledge and data we can connect the dots between oppressive systems and damaging outcomes. Those who are most impacted by social injustices such as racial disparity and inequity should also be part of the conversation to guide us toward the right questions and help us to avoid repeating old mistakes (104). They must be at the table in helping to develop the most innovative solutions. It is therefore critical that the GIScience community itself become more diverse, and work with changemakers (e.g., from local communities impacted by racial inequities) to round out the full picture of injustices at play and to accelerate positive change. In fact, local communities may even develop their own practical solutions that can be scaled to a level that government agencies can broadly use (e.g., refs. 105 and 106).

It is now possible to apply a social-justice lens more directly to maps and spatial analytics, to locate populations of concern, reveal and understand inequities in experiences and barriers to equality within a community, and support informed and equitable decision-making. The preparation of GIScience to accelerate social justice builds on a foundation of work aiming to improve decision-making and the well-being of all people and the planet. For example, an established application domain in GIScience is political redistricting, where the objective function is prescribed and “getting it right” is driven by concerns that all people, regardless of social capital, have voice (107, 108). Other well-established areas include mapping barriers to meeting life’s basic needs, such as accessing food (109) and green space (110). With the democratization of data (see next section), changemakers, with basic GIS skills, can build on the legacy of GIScience and accelerate social justice interests by leveraging multiple datasets and creating robust and comprehensive assessments for and with communities (e.g., see the various methods explained in ref. 111).

Given that GIScience has such close ties to industry, a growing number of companies in this space, both large and small, have launched substantive social-justice initiatives, leveraging significant internal resources as well as close partnerships with organizations such as the National Association for the Advancement of Colored People (112), the Dolores Huerta Foundation, the nonprofit design firm Territorial Empathy, and the W.K. Kellogg Foundation (WKKF). The vision shared by WKKF and the GIS company Esri, as one example, is to develop community-action hubs to inspire changemakers to lead conversations about racial inequity, foster collective understanding of the issues facing the community, equip these changemakers with the tools to develop solutions, and broaden and activate networks for implementation and positive change (113). In addition, on a broader industry front, Microsoft now has an Environmental Justice team with an emphasis on nature-based solutions for communities of color and various geographic approaches.

Also, while many science, technology, engineering, and mathematics (STEM) disciplines are aiming to diversify training programs, the breadth of applications of GIScience is a powerful aid in introducing diverse students to STEM; there are a variety of entry points to GIScience, and this creates access that is difficult to duplicate in other STEM fields. There are also diverse destinations for people with skills in GIScience, including all workforce sectors (government, industry, academia, and nonprofits) and many applied areas (transportation to health), and entry points are available for people graduating from a range of programs (technical program to PhDs). The variety of careers for people trained in GIScience means that training diverse students will have impact at scale (e.g., ref. 104). On the way to achieving this companies and organizations need to remain intentional in interning and hiring these students, investing in partnerships with Historically Black Colleges and Universities (HBCUs), Hispanic-Serving Institutions (HSIs), and Tribal Colleges and Universities (TCUs) along the way, and showcasing their diverse staff in conference plenaries, workshops, and demonstration sessions. Along this journey, it will be important to acknowledge the uneasy history with GI and racial injustice and address the barriers that have resulted in relatively few geography, cartography, and GIScience programs at minority-serving institutions.

Democratizing Access. Given the power of GIScience, there are equity implications for who has access to data, methods, software, and solutions. In the past two decades, GIScience has made significant gains in improving access to location-based methods through the creation of versions of software that serve diverse communities of users. In the early 2000s academics called for open-source GIS tools (e.g., ref. 114), and in response communities developed packages like QGIS (115) and R spatial...
By building communities of support for open-source GIS software, the field created no- to low-cost solutions that have improved access to GIS methods (e.g., ref. 117). However, open-source GIS tools have also been a powerful mechanism for building community through sharing, customizing, and iteratively building new solutions. Using open-source platforms, researchers have been able to implement and deploy cutting-edge methods that others can build upon to advance GIScience more rapidly. Open-source tools are also providing practical support for scientific reproducibility and replicability, as shared code aids documentation and enables the intricacies of implementation to be evaluated and repeated by others (118).

The continued rise of open source into the 21st century has brought further clarity as to what “open” should really mean. “Open” is used with regard to software and services in the sense of open source, but also open standards, open data, and open access have now become equally important, especially with regard to equity. Also, there is further realization, especially within the GIScience community, that “open” may not always mean “free.” To wit, although an organization may produce open-source code, it may still sell access to the solution built with that code or charge fees for the technical support and other services needed to maintain the solution. In the academic community, technical support for in-house, grassroots open-source projects often and suddenly disappears with the departure of the graduate student, postdoc, or faculty member who originally championed it (or when the grant money runs out).

As such, equally important is the clarification that a commercial, for-profit company can also be open by 1) creating, leveraging, and contributing to open-source coding projects; 2) providing avenues to directly read, import, and export data within its platform in hundreds of formats; 3) supporting the hundreds of geospatial standards currently in use (including those of the Open Geospatial Consortium and the International Organization for Standardization); 4) providing open (and often free) access to scores of application programming interfaces (APIs) and software development kits (SDKs); and even 5) providing its software, web solutions, and documentation support in many different languages, including those that read left to right and right to left, and use either a period or comma for a decimal. This is all with the investment in professional quality control and quality assurance to ensure that the software performs as advertised. This must also pervade across continued versions and security patches, and on a variety of operating systems, database management systems, open-source web servers, external software programs, and related technologies. Hence, we argue that the opposite of “open source” is not “commercial.” The opposite of “open source” is “proprietary.” Further, an open, interoperable, and standards-compliant platform strategy turns out to be not only an expression of corporate social responsibility but also a good business model (119).

Easy-to-use GIS software, that is professionally coded and supported, has been important in democratizing access to GIS methods. Just as statistical software based on a graphical user interface (GUI) created massive uptake of statistics (120), lowering knowledge barriers has created access. Indeed, access to these methods (also known as workflows) is emerging as yet another feature of openness (121) and is also an important ingredient in the replicability of GIScience (52). Commercial software platforms such as ArcGIS now also routinely include a number of free application templates and solutions that can be easily configured by the user. Continuous improvements to the user experience (UX) via the user interface (UX/UI), as well as to the performance of the software, exposes more capabilities while also simplifying the workflows that help researchers learn, use, share, and reuse methods and analyses. Distributed processing workflows enable reproducible GIS methods that leverage all manner of web services and other features of modern cloud and web browser technologies (122). GIS analytical methods can even leverage the simplest of building-block geoprocessing tools and preexisting web services that do not require any coding by the user. More advanced researchers can combine these simpler methods with task-specific analytical workflows that use specific programming routines built using custom Python or R libraries. All this, coupled with detailed documentation, a large catalog of low-cost or free online education classes and tutorials (including MOOCs), and online user support communities, is lowering the barrier of entry to global scientific research production, sharing, and replicability.

Increased access to spatial data is one of the most obvious ways in which GIScience is being democratized. Spatial data are more abundant, but with growing availability of low-cost or free datasets they are also more accessible. An important example of the benefits of making spatial data accessible is the USGS Landsat program, which has captured imagery since 1972 and switched from a fee-based to free-dissemination model in 2008 (123). Compared to 2009, in 2017 downloads of Landsat imagery had increased 20-fold to 2 million images (124). In fact, Esri downloads and processes hundreds of new Landsat scenes nightly and provides them in simple viewers on the web so that nonspecialists can better understand and use this valuable resource (125). Rather than just the picture or cached image of the Landsat scene, the view contains the full information content as a dynamic, high-performance image service that performs on-the-fly processing and dynamic mosaicking of Landsat’s multispectral and multitemporal imagery (e.g., ref. 126). Esri’s more dynamic tool allows the user to view the different bands, as well as the spectral and temporal profiles. While more people are able to use Landsat, the scale of studies has also increased (i.e., global mapping and monitoring of forests and water), enhancing societal benefits of this publicly funded data archive.

Other official data are being made more accessible, in part because web-based GIS has made it easier for research laboratories and agencies to create portals for data sharing. For example, it is common for cities to have open data portals, where data, typically generated for operations and planning, can be shared. In research laboratories, GIScientists are also using configurable web-application templates to make it easy for users of all abilities to make their own applications based on the foundational data provided by the laboratory, which will in turn lead to the use and adoption of those data, such as, for example, the use of imagery for burn-scar detection to create a baseline for forest regeneration and vegetation succession after a fire (e.g., ref. 127). Yet another way to increase the use and value of both data and information is via the aforementioned dashboards. These have an additional advantage in providing near-real-time displays for monitoring instruments, events, or activities, often times for “at-a-glance” decision-making, providing key insights on just a single screen.

Crowdsourcing, VGI (128), and participatory GIS are also mechanisms for democratizing creation and access to spatial data. Crowdsourcing and VGI are approaches to data collection that rely on people to contribute data based on their behavior or observations. These terms can be used synonymously, though typically crowdsourcing refers to a large number of people contributing data and VGI indicates data collected as map coordinates. Participatory GIS uses tools similar to VGI but usually includes some type of community empowerment.
Crowdsourcing, VGI, and participatory GIS approaches to the creation of spatial data provide an opportunity for nonexperts in GIS to engage in mapping, share knowledge, and create data. For example, OSM, arguably the most successful VGI or crowsource mapping project, has been used as a tool for community engagement and empowerment. As discussed in Data Representativeness, there are a growing number of businesses that generate commercial data products from mobility data crowdsourced from smartphones, often without full knowledge of the smartphone user. These datasets can be costly and can create data havens and have nots, based on financial resources, access to technology, or social capital. Users are becoming more aware of this practice by companies and educating themselves about shielding their personal data and movements. Activist employees in some of these companies are speaking out on the issue as well, seeking to change the practice and the company culture that engenders it (129).

One might consider the vast change in the quantities of spatial data a disruption in GIScience and more generally in society. As with any disruption, new opportunities are created for those most ready to pivot or take advantage of new circumstance. While there is a tendency to link greater data access to more equitable access, unfortunately access differentially empowers those that hold the knowledge and skills needed to utilize data. An unintended consequence of the disruption associated with spatial data availability is that it may put people lacking data science farther behind. A good analogy is high-performance computing (HPC), where the power to make use of the technology tends to widen the gap between the HPC-familiar and the rest. It is interesting to consider how the disruption of the 3Es coalesces with the disruptions associated with advancing technology and increasing spatial data, which will create technology as advances in data and technology have the potential to widen gaps yet the 3Es require new ways of operating to reduce gaps.

Discussion and Conclusion

The broad use of GIScience methods and data comes with a responsibility for leadership on the 3Es. As the creators of spatial data, methods, software, and a workforce that will infuse many disciplines and sectors, how we do our work and measure success matters. As we take the pulse of 3Es drivers and the response to them by GIScientists, we see encouraging signs of a community mobilizing for change. The list of professional development networks, professional societies, and individual initiatives mentioned at the beginning of the paper is a testament to progress in the 3Es and to note any potential for unintended consequences in the work.

Humility. Growing ethics, empathy, and equity in GIScience is a process that will require humility. We will get it wrong and make mistakes, and then we must commit to doing better. One of the authors (T.A.N.) has been working on research to reduce sampling bias in data (39). The research goal is to increase the representativeness of bicycling data by correcting bias in data generated by a fitness application, which oversamples bicycling patterns of men and of people aged 30 to 50 y. Data-correction models have been built in cities that had existing reference data, all of which have predominately White populations. In retrospect, the use of available training data has led to reinforcing systemic racism by building better data only in White communities. It can be painful to admit that our approach to research is part of the problem, but the reason we need to accelerate the 3Es is that the approaches we were trained in have left people behind.

Humility is required to admit we got it wrong, and courage is required to point out when others are designing or conducting research that reinforces exclusion. GIScientists, and ideally all researchers, must have the expectation that our colleagues will point out the potential of our work to reinforce systemic biases, and we must commit to doing the same for them. GIScience-related grants and journals can lead by explicitly requiring reviewers to evaluate papers in the context of the 3Es and to note any potential for unintended consequences in the work.

Broadening Measures of Excellence and Success. Focusing on equity will require that we broaden measures of excellence and success. In a recent graduate seminar taught to incoming students at the University of California Santa Barbara, during introductions nearly every student identified a hope that their work would positively impact the world beyond academia. The incoming students' values are a reminder that the emerging generation of talent will not be satisfied with careers based only on traditional measures of academic success. To attract the best talent and make space for the additional effort required to advance the 3Es, GIScience must grapple with how to value more diverse pathways. Contending with how to broaden measures of excellence is not unique to GIScience. As mentioned in the Introduction, NSF in 2010 began requiring the inclusion of a “Broader Impacts” section to grant proposals, signaling the value of socially impactful research. However, a barrier to broadening definitions of excellence is that impact can be hard to measure. At present, we often leave it to individuals to make value judgments about nontraditional measures of excellence and success. A starting point is to demonstrate that activities such as community engagement and knowledge translation have value, and this is happening. For example, the Canadian Institute for Health Research funds knowledge translation, or the conversion of basic research into actionable solutions, creating funding for students and researchers to connect with decision-makers. Funding knowledge translation also creates a workforce that is able to link basic research to decision making and creates a mechanism for valuing different types of academic work.

Diversify Your Network. Addressing pressing and current issues related to health, social injustice, and climate-change impacts requires creativity and diverse perspectives. Now is the time for all hands on deck. Both generalists and specialists from industry, government, and academia are greatly needed. Also needed are people trained in STEM and those deeply understanding cultural context. In other words, there is no time for silos in GIScience; more can be achieved and faster by working together. Interdisciplinary collaboration is not new in GIScience, but this article challenges GIScientists to accelerate solutions by regularly extending themselves to people working in diverse fields. If not having had the pleasure of asking a colleague to explain their jargon recently, it might be time to consider growing one’s network or to consider the diversity of that network. If a team is homogeneous in race, gender, or sexual identity, consideration should be given to how to grow that network to include more diverse collaborators. The time is past for waiting for the pipeline to diversify itself. There must be a more intentional building of teams that will support the training of more diverse people, which means seeking out diverse team members. In growing your network, it is important to be aware of the burden that can be placed on people when the same individuals are repeatedly asked to join teams. One may consider finding new people, by asking the handful of go-to people for recommendations. Buddy approaches, where the core is asked to invite another scholar to
Inclusive Education and Workflows. Creating multiple pathways into GIScience education can be one way to create inclusion in the field. Education in GIScience, as with many STEM fields, typically relies on a series where lower-level courses become prerequisites for later courses. When technical skills are required, it makes sense that knowledge builds. However, prerequisites can create barriers to inclusion. For example, while we delight in students graduating from high school knowing they want a career in GIScience, it is a relatively rare trajectory and students that discover the benefits of GIScience late can be challenged to catch up on prerequisites. Summer intensives, boot camps, and workshops in programming, statistics, remote sensing, and GIS can serve multiple benefits, including catching students up on prerequisites and creating more access to GIScience. We also need communities that know how to support diverse cultural contexts of students. Organizations such as the aforementioned NorthStar (https://gisnorthstar.org/), which supports and connects people of Black/African diaspora, are important groups to build partnerships and to grow inclusion and diversity in the field. In addition, designing workflows, where data management and analysis steps are clearly laid out, can allow broader utilization of advanced methods and spatial data. Workflows are a practical approach to reducing knowledge requirements needed to benefit from technical advancements and can broaden the equitable utilization of spatial data.

The field of GIScience must commit to accelerating the 3Es; it is the right thing for people and for the planet. It is the right thing for the culture of science and it will require fundamental rethinking of some of the traditional ways we operate. It is important for students, postdocs, and for serving as many segments of society as possible. Indeed, when we interact with emerging scholars and professionals in GIScience we are optimistic. A new generation is leaning into the 3Es. They are a generation that has grown up with knowledge at their fingertips and seem focused on how to leverage knowledge for impact. Amid rapidly advancing technologies and the ever-broadening reach of our field it is an exciting time to be a GIScientist, but continuing to attract talent will require programs, institutions, and industries committed to growing the 3Es, and opportunities for emerging GISScientists to contribute to a more ethical, empathetic, and equitable world.

Data Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. This paper was seeded by the presentation of T.A.N. in a webinar hosted by the Department of Geography at the University of California Santa Barbara entitled “Dangermond Perspectives on Emerging Trends in GIS” (https://www.geog.ucsb.edu/events/all/2021/webinar-dangermond-perspectives-emerging-trends-gis). We are thankful for comments from Fiona Goodchild and Aileen Buckley on earlier versions of the paper. We are also grateful to three anonymous reviewers and the chance to improve and tune our ideas by responding to their insightful comments.
12 of 12

https://doi.org/10.1373/pnas.2119967119

pnas