GeoJournal (2008) 72:239–244
DOI 10.1007/s10708-008-9190-4

Commentary: whither VGI?

Michael F. Goodchild

Published online: 26 July 2008
© Springer Science+Business Media B.V. 2008

For the most part, we find our way around the planet by tapping memory, perhaps by recalling past routes, or sometimes by accessing stored mental images that approximate maps. In other cases, however, we must rely on information collected and made available to us by others. People visiting strange cities need street maps to help them find their way around; planners need maps of wetlands and other constraints on development; pilots need air navigation charts; children need atlases and globes to learn about the Earth's almost infinite variety; and emergency responders need maps to plan systematic recovery efforts. Geographic information, whether in the form of maps, images, driving directions, or guidebooks is increasingly essential to many aspects of human existence.

As with many other forms of information, the process of acquiring, compiling, printing, and disseminating geographic information has been expensive. A topographic map costs the U.S. Geological Survey approximately $100,000 to produce, and publishers must make similar levels of investment in compiling and printing guidebooks. Mapmaking may involve the launching of satellites or leasing of aircraft; the use of photogrammetric machines costing more than $250,000 or elaborate surveying equipment; hiring and paying highly trained personnel to operate the equipment; and the expensive process of preparing material for printing. It follows that this process inevitably favored those geographic themes that were comparatively unchanging, allowing the information collected and disseminated to be valid for as long as possible; and favored those themes that had multiple applications, in order to maximize the potential interest in the product. Much of the traditional system of geographic information production, which tended to emphasize centralization in public-sector agencies, can thus be explained by simple arguments about scale economies and market size. Only central governments and a handful of publishers had the resources, mandate, and will to absorb the costs.

This world of national mapping agencies, standardized map series, and established publishing houses is the background against which the phenomenon of VGI must be understood. It explains why VGI is so interesting, provokes such a strong reaction in many quarters, and stimulates the kinds of writing represented in this special issue. The rapid growth of VGI in the past few years is one more step in a lengthy process that began almost two decades ago, and will likely continue for some time to come. It is one part of a fundamental transition as society redefines its vision of the role of public information in the early years of the 21st century.

This is an impressive collection of papers, and one of the first to tackle this expanding and timely topic.

M. F. Goodchild (✉)
Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA
e-mail: good@geog.ucsb.edu

It covers much ground, but inevitably there are gaps. In this commentary I would like to offer a few thoughts on two topics that are given less attention in the collection: first, the types of data that are suitably contributed by volunteers; and second, the question of society's right to open access to information. I end with some comments on the future of mapping and the long-term effects of the VGI phenomenon.

Types of VGI

As Flanagin and Metzger (this issue) make clear in their paper, the question of credibility has many dimensions. Mapping by national mapping agencies has involved high levels of expertise, in the form of cartographic skills, knowledge of the operation of machinery, and familiarity with the subject matter. The information produced by mapping agencies was and is credible first and foremost because its employees are perceived as experts by the user community, whose members are in turn willing to believe that map-making requires expertise. But these perceptions and beliefs have been shaken by the fundamental changes that have occurred in geographic information technologies in the past few decades. Map-making using the free tools provided by corporations such as Google clearly does not require great expertise in many cases, and users are encouraged in this belief by books such as *Mapping Hacks* (Erle et al. 2005) which have been written by authors with no obvious qualifications in cartography, and with no reference to the cartographic literature. These beliefs have become the basis for the rapidly growing community of *neo-geographers* (Turner 2006), who are convinced in the words of Rana and Joliveau (2007) that "the old geography involves a prescribed role/interaction between the four main components, namely the audience, the information, the presenter and the subject… In NeoGeography, there are however no such boundaries on roles, ownership, and interactions of these four components". In essence, the skills that cartographers working for mapping agencies used to rely on to design high-quality maps are now enshrined in software. Someone making a map using Google tools need make no decisions about fonts, symbolization, projections, or any of the host of details of cartographic practice, because all of those decisions were made when the software was designed. Only in the kinds of industrial-strength GIS represented by ESRI's ArcGIS does the user still retain substantial control over cartographic design, but even there most design decisions can be simplified by accepting defaults. Distributing software is in this sense a way of distributing expertise from experts to non-experts, but with the inevitable consequence that individuality is wrung out of the design process.

Over the centuries navigators needing to determine position on the Earth's surface have relied on a variety of measurement techniques, some of them very sophisticated in their use of instruments and arithmetic calculations. The aircraft navigator of World War II, trying to locate a night-time target in Germany by sighting stars through the cockpit window, needed very substantial training and skill to be successful. Today, of course, determination of position has become trivially easy with the adoption of GPS, whether it be by sailors navigating over water, or drivers navigating through cities, or hikers navigating through wilderness. The skills that used to be needed by topographers are in many ways now redundant, once again leading to the widespread belief that anyone can be a map-maker—that there is no longer a distinction between expert and non-expert where geographic data acquisition is concerned.

The question of expertise regarding subject matter is more problematic, and provides a basis for differentiating between data types. Take first the example of placenames, and the construction of gazetteers (defined as collections of triples <name, location, type> where type is one of a controlled set of geographic feature types; Hill 2006). Placenames cannot be observed from space or from the air, and are in many cases not displayed in any form on the geographic landscape, so map-makers have traditionally consulted with local residents to identify the names attached to features. That said, however, local usage is often ignored or overridden when it duplicates, or when it is too informal or politically incorrect (Monmonier 2006), and a hierarchy of geographic names authorities exists in many countries to resolve ambiguities and conflicts.

Today the need to standardize placenames seems less compelling, since information technology is capable of providing automatic translation between equivalent terms. Why not allow gazetteers to capture all of the formal and informal names attached to features, particularly when their purpose is to aid users in finding places or in decoding references to them in speech or text? In this sense, then, VGI projects to collect names and descriptions of features, such as Wikimapia, seem particularly appropriate, and likely to lead to much richer gazetteers than has been possible in the past. There is no argument here that compiling such information requires advanced subject-matter expertise, since placenames are part of everyday linguistic usage.

Nevertheless projects such as Wikimapia are open to questions about quality. When someone creates a Wikimapia entry there is no immediate consultation with neighbors or others in the local area for confirmation, as there was in traditional topographic mapping. Yet Wikimapia is able to exploit the same kind of error-correction process used by Wikipedia: entries are to some extent monitored by volunteers, and entries are open to editing by anyone who disagrees with them. The consensus appears to be that this process works reasonably well with Wikipedia (Giles 2005), especially with prominent or important entries. Moreover, residents of a neighborhood are inherently experts in the local area. So for Wikimapia nearby residents provide a potential source of expertise about every entry, whereas for the more obscure Wikipedia entries there may be no one capable of detecting errors.

OpenStreetMap provides a rather different example of a map of streets, railroads, and rivers compiled by volunteer effort. As in the gazetteer case, there seems to be no need for any special expertise, since anyone is capable of following and naming streets and other simple features in the real world, and managing the data created with a GPS tracking device. Here the entries created by volunteers must necessarily agree with neighboring entries—the streets, rivers, and railroads acquired by one volunteer must edgematch with those collected by others, and names must be consistent. Thus the task of compiling independently contributed pieces of the patchwork necessarily imposes some degree of quality control. One might term this process *structured*, to distinguish it from the essentially unstructured process by which entries in Wikimapia and other VGI gazetteers are compiled.

In other cases, however, tradition demands that maps be made by experts in the subject matter. Take the case of soil mapping, which has been performed by scientists trained in the nature and distribution of soils. Against this tradition, it would be foolhardy to suggest that a non-expert could be permitted to collect and publish soil data, or that the result would be credible to the farmers and geotechnologists who make use of soil maps.

But changes have been occurring in the world of soil mapping that are parallel and in many ways comparable to changes in other parts of the world of geographic information. Farmers who have invested in the tools of precision agriculture are now able to collect and manage soil data that are far more precise and detailed than traditional soil maps, and far more useful for day-to-day decision making. Moreover these data address parameters of direct interest to farming, such as soil moisture content and soil pH, that are only indirectly addressed by the classification systems that are the primary subject of traditional soil mapping; and measurement of soil moisture content and soil pH using appropriate instrumentation is well within the capabilities of someone without advanced training in soil mapping. Why, then, is it necessary to have advanced training to make a soil map, but not to measure the properties needed for a specific application of soil mapping?

Perhaps the answer lies in the scale economies that were the subject of the opening comments above. To justify the high costs of making soil maps, it was necessary to design them for as many purposes as possible. Rather than measure specific parameters of soils, therefore, mapping focused on broad classes that captured as much of the general variation as possible, by minimizing aggregate within-class variation and maximizing aggregate between-class variation in the many measurable properties, but without necessarily doing so for any *specific* property. The result was a compromise, since some degree of precision was lost in the representation of any specific property—and it required that soil mapping be conducted by experts, since only they could make the complex judgments necessary. Once those scale economies disappeared, however, as the high cost of entry into the soil mapping business shrank almost to zero and it was no longer necessary to map for all possible applications, so too did the need for experts in the soil-mapping process.

It seems, then, that all three arguments for mapping expertise presented at the beginning of this section—the need for cartographic skills, skills in the operation of complex measuring instruments, and familiarity with the subject matter of mapping—may have disappeared as a result of technological change, and not only for conceptual simple types of geographic information such as placenames but for the most complex types such as soils. Mapping expertise may no longer be one basis on which to judge credibility, or to distinguish the expert from the non-expert.

Open access to information

Society has adopted a complex set of procedures to ensure that citizens have appropriately constrained access to information. Information that would compromise national security if placed in the wrong hands is protected by *classification*, for example, and information on individuals that might be perceived as invading privacy is protected by removal of identifying information, and in some cases by spatial aggregation. Nevertheless the collective amount of personal information held by government agencies and private corporations is vast, and its security is periodically compromised by leaks, theft, and other kinds of disclosure. One result is *identity theft*, a form of crime that depends on its perpetrators being able to assume the identities of its victims. Curry (1998) has written about the particular forms of privacy invasion that are made possible by detailed geographic information.

Other reasons besides national security and protection of privacy can also be used to justify denying certain kinds of information to the general public. Consider the following example, based on incidents that occurred in Santa Barbara in 2007. It is now widely accepted that global climate change will lead to rising sea levels and to consequent problems for coastal communities. The Light Blue Line project (http://www.lightblueline.org) was devised as a way of drawing attention to predictions about sea level rise, and thus to the dangers of global climate change and greenhouse-gas emissions: a line representing the predicted position of the coastline after a 7 m rise would be painted on the streets and sidewalks of the city. A grant was obtained, and approval was given by the city's administration. But a group of wealthy landowners and developers objected, threatening legal action on the grounds that the painted line would adversely affect the value of properties on its downhill side. Approvals were reversed and the immediate plans were abandoned.

One can scarcely object to the existence of scientific forecasts in this case, since they are widely known—rather, the problem appears to lie with the way the information was to have been presented, and the fact that it would be available to all in readily understandable form. Forecasts of future sea level positions can easily be found on the Internet, in the form of mashups with Google Earth or Google Maps, and these are accessible to anyone. But some invisible threshold seems to have been passed by the Light Blue Line project, in proposing to paint the line on the Earth itself rather than in cyberspace.

During an emergency evacuation it is critical that accurate information be available about people who are insufficiently mobile—the *shut-ins*. At first sight VGI would seem to be an excellent mechanism for collecting and maintaining such information, since it could be provided by the shut-ins themselves, or by concerned neighbors, and regularly updated. But such information is also invaluable to criminals, and clearly must be treated as highly confidential. During actual emergencies there is always pressure to relax restrictions on access to geographic information—for example, to persuade the private-sector owners of detailed information on electrical networks to allow that information to be distributed temporarily to emergency managers (National Research Council 2007). Later, however, it may be difficult to reimpose the normal constraints by recalling or destroying all distributed copies.

Citizens volunteering geographic information clearly do so in the belief that it will be open, accessible, and free, and may even be of critical value in some future emergency. The OpenStreetMap, for example, might well prove valuable in routing emergency vehicles in the aftermath of some disaster. But it is also easy to imagine circumstances in which volunteered information, about oneself or one's neighbors, leads to forms of exploitation that were difficult to foresee. Research is clearly needed to define the limits of VGI in this context.

The long term

Like many features of Web 2.0, VGI has all of the aspects of a fad, a flash in the cartographic pan. It has clearly captured the imagination of a significant fraction of the general public, at least on the favored side of the digital divide, and is absorbing a substantial amount of volunteered time. But its rhetoric—"Let's describe the whole world" (Wikimapia), "a map of the whole world…made by people like you" (OpenStreetMap)—implies a one-time effort rather than a continuing process of update to keep track of a constantly changing world. Will the enthusiasm for maintenance be as great as the initial enthusiasm, or will these projects start to resemble the fate of the USGS's National Topographic Series, steadily growing more out of date? And who will preserve the results, to ensure that they are still available decades in the future?

I began this commentary with a reference to the broader transitions in the world of geographic information of which VGI is just one part. With the collapse of the scale economies that forced a centralization of mapping, we are seeing a steady trend to localization, through increased local production and consumption of geographic information. Arguably this is the natural state, reflecting the local nature of geography and the role of local expertise. In future we may even come to think it odd that in the 20th century it was customary to maintain a large establishment of expertly trained cartographers in Washington and other centers, and to fly them to different parts of the world to collect and compile geographic information, rather than relying on local experts.

There are many distinct types of geographic information, and several taxonomies have been devised for various purposes. It is customary to define *framework* data as consisting of those types that can be used to georeference other information—a street map, for example, can be used to georeference information by street address, and a gazetteer to georeference by placename. Hydrography is used as a framework by water resource specialists, and elevation by geomorphologists; and geodetic data provide the most accurate framework. Framework data were seen as the most valuable component of the National Spatial Data Infrastructure, since they made so much other activity possible, and was given highest priority in planning and coordination. But framework data are now available from numerous sources, and direct georeferencing using GPS makes them redundant in many cases.

In the rapidly evolving world of geographic information a more effective distinction may be emerging that is of particular importance to VGI. Some types of geographic information can be sensed remotely, allowing satellites and airborne platforms to collect vast amounts of cheap data without accessing the ground or its inhabitants. Different parts of the electromagnetic spectrum can be used to detect a wide range of phenomena, and active sensing using light, radar, and microwave further extends the options. Economies of scale are critical, forcing data collection to be organized centrally.

But other types of information are not detectable remotely, and require local consultation. This is the arena of VGI, and of the rise of local expertise to replace centralization. In the long term, then, it seems that the world of geographic information will bifurcate even more than at present. Citizens will be directly involved in the creation of those kinds of information that cannot be sensed remotely, through locally defined activities. New kinds of information that directly address specific applications will become more available, replacing such generically motivated products as soil maps. Citizens are particularly suited to identifying errors, and to providing rapid notification of change. In the world of local mapmaking, then, distinctions between expert and non-expert, between producer and consumer, will steadily disappear.

Where does this leave the geographer, the traditional custodian of expertise? To the general public, and neogeographers in particular, map-making has become something that anyone can do, and if that is what geographers did then we are all now potentially geographers. There is an urgent need to explain why geography is more than making maps, and why the methods of a geographer are powerful and far from intuitive. The literature of geography represents a vast and steadily growing collection of knowledge about how the world looks and how it works, obtained over decades through theory-building and the careful analysis of geographic information. Yet this knowledge remains largely hidden to the average citizen, who sees instead the intuitive simplicity of VGI. We need to find ways to project geographic knowledge into this world, making it more accessible to the average citizen. For example, a platform such as Google Earth could be used to show the dynamics of geographic phenomena and simulations of Earth futures, both important parts of the Gore Digital Earth vision (http://www.isde5.org/al_gore_speech.htm). Instead of the current free-for-all of geographic facts collected by sites such as Wikimapia, citizens could be invited to provide specific kinds of information of greater relevance to geographic understanding, in the spirit of citizen science—"Let's *understand* the whole world".

References

Curry, M. R. (1998). *Digital places: Living with geographic information technologies*. New York: Routledge.

Erle, S., Gibson, R., & Walsh, J. (2005). *Mapping hacks: Tips and tools for electronic cartography*. Sebastopol, CA: O'Reilly.

Flanagin, A. J., & Metzger, M. J. (2008, this issue). The credibility of volunteered geographic information. *GeoJournal*.

Giles, J. (2005). Special report: Internet encyclopaedias go head to head. *Nature, 438*, 900–901.

Hill, L. L. (2006). *Georeferencing: The geographic associations of information*. Cambridge, MA: MIT Press.

Monmonier, M. S. (2006). *From squaw tit to whorehouse meadow: How maps name, claim, and inflame*. Chicago: University of Chicago Press.

National Research Council. (2007). *Successful response starts with a map: Improving geospatial support for disaster management*. Washington, DC: National Academy Press. http://books.nap.edu/catalog.php?record_id=11793

Rana, S. & Joliveau, T. (2007). Call for papers: Special issue of the Journal of Location Based Services on Neogeography

Turner, A. (2006). *Introduction to neogeography*. Sebastopol, CA: O'Reilly.