GeoJournal (2008) 72:137–148
DOI 10.1007/s10708-008-9188-y

The credibility of volunteered geographic information

Andrew J. Flanagin · Miriam J. Metzger

Published online: 24 July 2008

Abstract The proliferation of information sources as a result of networked computers and other interconnected devices has prompted significant changes in the amount, availability, and nature of geographic information. Among the more significant changes is the increasing amount of readily available *volunteered* geographic information. Although volunteered information has fundamentally enhanced geographic data, it has also prompted concerns with regard to its quality, reliability, and overall value. This essay situates these concerns as issues of information and source *credibility* by (a) examining the information environment fostering collective information contribution, (b) exploring the environment of information abundance, examining credibility and related notions within this environment, and leveraging extant research findings to understand user-generated geographic information, (c) articulating strategies to discern the credibility of volunteered geographic information (VGI), including relevant tools useful in this endeavor, and (d) outlining specific research questions germane to VGI and credibility.

A. J. Flanagin (✉) · M. J. Metzger
Department of Communication, University of California, Santa Barbara, Santa Barbara, CA 93106, USA
e-mail: flanagin@comm.ucsb.edu

M. J. Metzger
e-mail: metzger@comm.ucsb.edu

Keywords Volunteered geographic information · VGI · Credibility · User-generated content · Social computing · Information trust

The contemporary media environment is remarkable in its capacity to promote, support, and sustain collective endeavors among disaggregated individuals. By its very design, the Internet positions "intelligence" (discrimination and processing functions) primarily at the periphery of the network, in the hands of individuals. One function of this basic design decision is that interactivity is literally built into the Internet through a dynamic system where users play roles of both information consumer and information provider. As a consequence, this structure is particularly well suited to *collaboration* among individuals.

Indeed, in the brief history of the Internet there has been a steady appreciation of the hallmark of digital media: the ability to connect to others and share information, easily, quickly, and across distance and time. This appreciation has bloomed in parallel to significant technical developments in both access devices (e.g., computer technologies becoming simultaneously more powerful and less expensive) and infrastructure improvements (e.g., the wide availability of broadband connections). The end result has been the recognition of the power of the Internet and related technologies to take full advantage of their enormous

scale in ways that capitalize on the diversity and knowledge of multiple users.

Spatial and geographic applications of this basic capacity abound. Examples range from worldwide mapping and map annotation tools to environmental design endeavors that leverage the esoteric and situated experiences of locals for their success. These and many other such examples suggest the enormous knowledge assets that reside in collectives, which until recently remained largely untapped due to insurmountable coordination costs. On the one hand, these resources represent significant new benefits to the study of geography. On the other hand, they also suggest serious threats to the veracity of geodata, and the degree to which information thus provided can and should be trusted.

This essay considers these issues by examining the credibility of volunteered geographic information (VGI). First, the social and information environment fostering collective information contribution is investigated, with an eye to specific geographic applications. Next, VGI is situated in terms of information and source *credibility*. To do so, the environment of information abundance is explored, issues of information and source credibility are located among other related terms and within their specific research heritage, and issues particularly germane to geographic data are considered. In the process, existing research findings are leveraged and applied to user-generated geographic information, in order to learn from relevant past work. Finally, some prototypes for developing tools that might be useful in this endeavor are considered, and research directions and conclusions regarding how best to discern the credibility of VGI are offered.

Social computing and volunteered geographic information

The proliferation of social computing practices has dramatically increased the amount of user-generated content available online. Recent data show that 35% of Internet users in the U.S. have created content and posted it online; 26–34% have shared something online that they created themselves (e.g., photos or videos); 32% have rated a product, service, or person using an online rating system; 20% have created a personal profile that others can see in social networking sites or elsewhere; and 8% keep blogs, which are read by 39% of Internet users in the U.S. (Lenhart 2006). When coupled with the fact that over 70% of all adults in the U.S. are now online regularly (Harris Interactive 2007; Lenhart 2006), these figures suggest that a large proportion of those in the U.S. are participating in some meaningful way in the creation of their own information environments. Inasmuch as web-based information constitutes a meaningful component of the information environment more broadly, even those not directly participating in or contributing to this information repository are potentially affected by its existence.

This trend toward user-generated content has had profound impacts on geographic information that is provided, altered, or shared online. For example, the online map and satellite imaging resource Wikimapia (http://wikimapia.org) currently offers nearly 7 million places that have been identified and annotated by its users, thus making nontrivial inroads toward its goal to "describe the whole planet Earth." Google Earth similarly provides opportunities for users to offer location-based data, by annotating specific reference points with photographs and textual descriptions. In its first 2 years, Google Earth has attracted 200 million users (Google 2007).

The release to the public of the Google Maps application programming interface (API) has resulted in a further wellspring of user-generated resources. Use of the API to create "mashups," or web applications that combine data from multiple sources to form a new integrated resource, has resulted in a wide variety of geographic tools relying on user-generated content (Miller 2006). For example, Google Maps have been leveraged to identify and share hiking or biking routes, to locate and share anomalies or points of interest discovered in existing Google Earth maps, to annotate and share local neighborhood or real estate information, and to combine specific location information with items offered for sale on Craigslist, to name only a few current outcomes. Given that many such applications have been developed by for-profit entities, there is often a vested interest in popularizing these tools, which in turn enhances greater participation on the part of individual users.

Moreover, pursuits that are ordinarily the province of geography, as well as many more typically outside of its domain, are implicated. OpenStreetMap (http://www.openstreetmap.org), for instance, relies on individuals' GPS data as they traverse specific areas, in

order to create comprehensive and up-to-date maps. Individuals are in this instance used as carriers of sensors to measure their local environments (Goodchild 2007). After sufficient data points are compiled, users can view, edit, and use resultant maps, in many cases in creative ways such as rendering 3D maps or adding ski run data to mountain areas. Although not originally intended as providing geodata, the photo-sharing site Flickr (http://flickr.com/) also contributes tremendous amounts of place-based information via its members' shared photographs of specific locations. As photographic information accumulates, images on Flickr can constitute a relatively rich location-specific resource.

In these and other ways, geographic data are increasingly provided by individuals, who in most cases are not trained or even necessarily interested in geography as a science. This marks a trend toward increasing publicly volunteered geographic information, where various forms of geodata are provided voluntarily by individuals (Goodchild 2007). Depending on its instantiation, VGI can be viewed either as an extension of public participation geographic information systems (PPGIS; Sieber 2006), collaborative GIS (Balram and Dragicevic 2006), participatory GIS (Elwood 2006b; Kyem 2004), and Community Integrated GIS (CIGIS; Elmes et al. 2004) to laypeople, or as a version of these that is more inclusive by virtue of its emergence among end-users, based on their own particular interests and goals (sometimes called "GIS/2;" see Miller 2006). Indeed, some argue that contemporary geographic and spatial practices such as those described above signal *neogeography* (Jackson 2006; Turner 2006), based on nontraditional GIS techniques that spawn "geography without geographers" (Sui 2008, p. 5).

The availability of user-generated geographic data has undeniably improved geographic information in several ways. Tremendous amounts of place-based data, images, and other geographically relevant information are now readily available, greatly enhancing the overall body of environmental knowledge. Moreover, individuals are in many cases in the best position to provide information that requires indigenous experience, esoteric understanding of a particular physical environment, and current information about local conditions. Individually experienced and contributed information has natural advantages under such conditions.

However, VGI also raises concerns with regard to its quality, reliability, and value as an information resource. For example, the multiplicity of sources that ensure vast information availability also make assessing the credibility of information extremely complex. Moreover, the origin of geographic information, and thus its quality and veracity, are now in many cases less clear than ever before, resulting in an unparalleled burden on individuals to locate appropriate information and assess its meaning and relevance accurately. Doing so is highly consequential: assessing credibility inaccurately can have serious scientific, social, personal, educational, and even political consequences. As a result, determining the credibility of information becomes critical as people process VGI gleaned from digital media. To fully comprehend these issues, the technical and social environment of VGI must be understood, as discussed next.

The environment of volunteered geographic information and implications for information and source credibility

Although there is some disagreement on its definition, the proliferation of user-generated content is viewed as a by-product of the so-called "web 2.0" environment (see O'Reilly 2005), which typically denotes an atmosphere in which network technologies are used to leverage the potential contributions of a wide variety of users, each of whom may contribute value in some manner to collective endeavors. In contrast to a static information delivery platform, this view conceives of the Internet as a dynamic collaborative environment in which diverse information, opinions, experiences, and skills can be aggregated to provide substantial resources. The essential premise is that given efficient means of information sharing and participation, collective benefits will emerge from aggregated individual contributions.

Information abundance, gatekeeping, and quality control

Perhaps the greatest change to emerge from web 2.0 tools and practices is that digital media have provided access to an unprecedented amount of

information available for public consumption. Until recently, the enormous cost and complexity involved in producing and disseminating information—particularly that of a scientific nature—limited the number of information providers, who generally had substantial investment in either the information itself or in the apparatus required to deliver it. Digital network technologies, however, have lowered the cost of information production and dissemination, thus increasing the sheer amount of information and the number of information sources available.

In tandem with information abundance come special problems for determining information and source credibility. Several factors contribute to this, including difficulty in locating and authenticating digital information sources (see, for example, Bose and Frew 2005; Frew 2007); decreased costs of information dissemination coupled with a lack of professional gatekeepers and quality control standards; and the lack of user familiarity with new forms and genres of information available online, among other issues. With regard to amateur mapping applications, Tulloch (2007a) notes "Many of these new applications are developed by individuals with no background or interest in the academic traditions associated, at the very least by implication, with their efforts" (n.p.). The combination of the vast quantity of and accessibility to digitally stored and transmitted information has prompted concerns about its credibility because, as Rieh and Danielson (2007) argue, this combination creates greater uncertainty regarding both who is responsible for information and, consequently, whether it can be believed.

Two particularly important issues in this regard are the nature of gatekeeping in the digital media environment and the level of ambiguity surrounding both the source and the context of information. Scholars in communication and information science have pointed out that information posted on the web may not be subject to filtering through professional gatekeepers and, as a result, digital information may be more prone to being poorly organized, out of date, incomplete, or inaccurate (Flanagin and Metzger 2000, 2007; Metzger et al. 2003; Rieh and Danielson 2007). Others have noted that digital media sources sometimes lack traditional authority indicators such as author identity or an established reputation (Fritch and Cromwell 2001; Metzger 2007).

Yet, source information is often crucial to credibility because it is the primary basis upon which credibility judgments are thought to rest. At the same time, however, there are several ways in which the source of information is problematic in the digital media environment. In some cases, source information is unavailable, masked, or entirely missing from a web site, mapping application, wiki, and so on. In other cases, source information is provided, yet hard to interpret, such as when information is co-produced, re-purposed from one site, channel, or application to another, or when information aggregators display information from multiple sources in a centralized location that may itself be (inaccurately) perceived as the source. These technological features create a kind of "context deficit" for digital information (see Eysenbach 2008) that makes interpreting its credibility difficult. Moreover, the hyperlinked structure of the web contributes to this deficit by making it psychologically challenging for users to follow and evaluate various sources as they navigate quickly across multiple sites online (Eysenbach and Kohler 2002).

These concerns are reflected by Callister (2000) who argues that standard conventions of determining credibility break down in cyberspace. Traditional solutions to credibility assessment include granting credibility to some representative believed to provide reliable information (e.g., a government agency such as the U.S. Geological Survey) or granting it by credential (e.g., to a professor or research scientist in a particular field). This works, he says, only when there is a limited number of sources and when there are high barriers for access to public dissemination of information, since these conditions create a meritocratic filtering process where only those with something of merit to say are published (p. 412). In other words, these conditions create an environment of information scarcity where it is possible for gatekeepers to produce and filter much of the information available, and where those gatekeepers have incentive to uphold quality standards. The Internet presents a very different environment—one of information abundance—which makes traditional models of gatekeeper oversight untenable due to the sheer volume of information that would have to be vetted.

Information and source credibility

These issues all revolve in various ways around the notion of information and source *credibility*. Although there is no one, clear definition of credibility, it is generally thought to be the *believability* of a source or message, which is composed of two primary dimensions: trustworthiness and expertise (Hovland et al. 1953). While trust and expertise have meaning separate from credibility and from each other, credibility is usually conceived as possessing at least some degree of both trust and expertise in combination. So, one can trust someone who has no expertise (e.g., someone you feel is a good person but doesn't know anything about a particular issue), and one can perceive someone to be an expert but not trust that person (e.g., an elected official). Additionally, trustworthiness and expertise have both objective and subjective components. That is, trustworthiness is a receiver judgment based primarily on subjective factors. Expertise can also be subjectively perceived but includes relatively objective characteristics of the source or message as well (e.g., source credentials or information accuracy; see Metzger et al. 2003; Tseng and Fogg 1999 for extended discussions of these points). Thus, credibility is a complex concept that revolves around the believability of some source or information based on notions of its trustworthiness and expertise.

The study of credibility is highly interdisciplinary, and definitions are also somewhat field-specific. For example, many scientific communities, including information science perspectives, view credibility as an *objective property* of information "quality," or the degree to which information can be considered to be accurate, as judged by accepted standards or by experts in a particular domain. By contrast, the fields of communication and social psychology treat credibility more as a perceptual variable: credibility is not an objective property of a source or a piece of information, instead, it is a *subjective perception* on the part of the information receiver (Fogg and Tseng 1999; Gunther 1992). As such, the credibility of the same source or piece of information may be judged differently by different people. This suggests a distinction between the "accuracy" of information and its "credibility." While accurate information in most cases is likely to be perceived as credible, technically inaccurate information can also be perceived as credible as long as the information consumer believes it. This applies to both objective, factual information (e.g., a patient believing that a discredited medical treatment will work) as well as more subjective information (e.g., believing others' statements about a political candidate's positions).[1]

Conceiving of credibility as a perceptual variable highlights the fact that trust and expertise are problematic terms. For example, information science perspectives that view credible information as only that which is "accurate" lean too heavily on expertise: nonexperts can also be credible, and many studies have found instances where local knowledge or expertise has eclipsed that of credentialed experts (see, for example, Fischer 2000). Experts are also wrong sometimes, even though they may be trusted. So, who is an expert and whether those experts earn the trust of others can be separated from the accuracy of the information they provide. Perceptual conceptualizations of credibility, such as those found in communication and social psychology, better allow for these possibilities.

Viewing credibility as a perceptual variable accommodates the core issues of most publicly volunteered geographic information. While credibility-as-accuracy is adequate for describing and evaluating scientific knowledge production, credibility-as-perception enables a better understanding of knowledge production that is collaborative, distributed, collective, and conflictual, because it highlights, and makes central, the notion of *believability* (or what one finds to be credible) as an appropriate concern. Indeed, many forms of VGI today are less about scientific data accuracy and more about which information, opinion, or perspective people believe (e.g., which highway lookout point is best to view the Grand Canyon or by whose map electoral districts should be drawn). So, while credibility-as-accuracy is an appropriate concept for those who have a "factual" relationship with geospatial information (as do most scientists), credibility-as-perception is more useful for those who use VGI for social,

[1] The notion of credibility is allied closely with several concepts, including trust, reliability, accuracy, reputation, authority, and competence. Although several of these concepts include both of the core dimensions of credibility, some more closely resemble the trustworthiness dimension (e.g., reputation, reliability, trust), while others tend toward the expertise dimension (e.g., quality, accuracy, authority, competence).

communal, or political purposes (see Miller 2006 for a discussion of these different constituencies for GIS information). In this way, then, credibility-as-perception is critical to understand the social and political power of VGI.

Credibility of volunteered geographic information

Traditionally, a small number of central, usually professional information sources such as relevant governmental agencies, cartographers, geographers, or other entities endeavored to provide relatively objective geographic data in an authoritative manner. In such cases, credibility was granted based on the perceived authority of these few entities, which was generally agreed upon within the relevant scientific community and was (dis)proven over time. To maintain their credibility, these sources had high incentive to uphold information quality standards. Within an environment of information scarcity, they functionally served as the credible information sources, given the high costs of information contribution and dissemination.

Such sources were, and continue to be, held to the quality standards of the scientific community in which their expertise is earned. Thus, when a source such as the U.S. Geological Survey (USGS) endeavors to provide to the public maps of the world and makes mistakes in this undertaking, their long-term credibility may be called into question, just as would be that of any authority. Consistent errors, of course, erode trust in any source over time.

With regard to more clearly volunteered geographic information—that is, information contributed by members of the public who are not geographers or even scientists—it is useful to consider credibility according to the degree to which people's spatial or geographic information is unique and situated, and the extent to which its acquisition requires specialized, formal training. For example, several instances of VGI involve perceptual information that can only be reliably known and communicated by "locals" who are immersed in their environments in various ways.[2] As mentioned earlier, such information is most closely allied with views arguing that credibility is properly rooted in *perceptions* about the relative trustworthiness and believability of information or sources (rather than in its accuracy).

Knowledge about perceived landscape boundaries (Weiner and Harris 2003); community-based maps and tools used to enhance participatory urban planning or to reclaim territories (Peluso 1995; the Mumbai Free Map, http://mumbai.freemap.in/); information about housing, crime, and current land use from urban neighborhood residents (Elwood 2006a); and perceptual data on perceived spheres of geolocational influence (e.g., with the CommonCensus map project, http://www.commoncensus.org) all represent instances of situated and unique knowledge that do not require formal training, which are being harnessed using contemporary network applications. In such cases, no one source is objectively right or wrong, and it does not make sense to assess information based on some objective notion of information "quality" or "accuracy." Nor does information credibility rest on traditional "authorities," who hold the types of credentials typically associated with the science of geography. Rather, credibility rests on the extent to which a representative sample of people provide their personal input honestly and accurately.

In fact, in such instances the ability to aggregate information and to connect individuals to one another provides new potential for *undermining* traditional authorities, and network methods for determining information credibility (Flanagin and Metzger 2008). The potential for peer-to-peer credibility assessment, for example, can be seen in numerous applications, ranging from people's corrections of inaccurate Wikimapia entries to blogs and other forums where individuals pool firsthand experiences to create and maintain accurate spatial information. Paradoxically, then, while digital media and information abundance may complicate people's confidence in and knowledge of who is an authority, electronic networks and social computing applications make it easier to harness collective intelligence to assess and evaluate information and sources online. Social computing tools and applications can thus "replace the authoritative heft of

[2] Instances where volunteered geographic information does not involve knowledge that is exclusive to "locals" immersed in their environments include tourists' commentaries, images, or maps of a geographical location; or non-local "experts" volunteering opinions about a local resource management controversy or land use plan.

traditional institutions with the surging wisdom of crowds" (Madden and Fox 2006, p. 2).

In essence, digital media enable the uncoupling of credibility and authority in a way never before possible by calling into question conceptions of authority as centralized, impenetrable, and singularly accurate. The result may be a shift from a model of single authority based on information scarcity and hierarchy to a model of multiple distributed authorities based on information abundance and networks of individuals. Indeed, bottom-up assessments of information quality constructed through collective or community efforts (e.g., wikis, ratings and reputation systems, or social networking applications) may in many cases be emerging as new arbiters of credibility.[3]

That said, in situations where individuals' volunteered spatial or geographic information is less unique and not situated within a specific context in any important way, the grounds for credibility assessment are significantly shifted. When information can reliably be judged from a presumably objective standpoint, based on commonly shared and widely held standards among those possessing some expertise, assessing credibility in terms of its accuracy and assessing sources in terms of their credentials may be warranted. Such information is most closely allied with views arguing that credibility is properly a function of the relatively *objective properties* of information, rather than a matter of individual perception.

In such cases, the objective characteristics of geographical information, although always open to debate within a scientific community, functionally limit discussions of credibility to only those types of information for which there is some objective standard of "truth" or "fact." Under these circumstances, the accuracy of data objects is particularly critical since both upstream and downstream use can be tainted by inaccuracies (Frew 2007). For example, people's annotations of places or of photographs linked to physical artifacts contribute to the locus of knowledge about geographic locations, yet they are properly evaluated by virtue of their accuracy and reliability across observations (where relevant) or proximity to expert knowledge, rather than by their representativeness within a population. Although such information is still volunteered by individuals, its credibility should be tied to widely shared standards of accuracy or quality within relevant lay or expert populations. Landmarks whose locations are misattributed by Wikimapia users, for instance, do in fact constitute errors, which impact the credibility of the information as well as the source's perceived credibility.

Situated between the extremes of high versus low information uniqueness, situatedness, and formalized training are hybrid efforts, where there is some degree of short-term expertise involved, which is typically *learned* by participants, or that relies in large part on specialized interest or skill. For example, the Audobon Society's Christmas Bird Count (see http://www.audubon.org/bird/cbc/index.html) has for 108 years marshaled the resources of volunteers to observe and report on early-winter bird populations. In 2007, more than 50,000 observers provided data on the existence of specific bird species by location, thus providing an unparalleled database of migratory and bird population data. Similarly, the Center for Remote Sensing and Spatial Analysis, working with the New Jersey Department of Environmental Protection, has enlisted the aid of volunteers to survey and map New Jersey's existing vernal pools, in order to certify them to receive regulatory protection as critical habitat areas (http://www.dbcrssa.rutgers.edu/ims/vernal/; see Tulloch 2007b for more information). In this case, volunteers receive specialized training in order to ensure validity and reliability, and data are pooled to create interactive maps displaying their results.

In such cases, information is still provided by several volunteers, but only after specific training is undertaken in order to increase accuracy and

[3] Although this new credibility model offers exciting promise for information evaluation, it can also be problematic. Collective intelligence can function well in many circumstances, yet it is also subject to biases through processes of bandwagon effects and groupthink. Stated differently, crowds may not always be so wise: group opinion conformity can result from collective deliberation because individuals' judgments are often influenced by others' judgments. So, if biases are introduced early in the deliberative process, group dynamics such as the tendency toward social reaffirmation may end up perpetuating rather than challenging these biases. This in turn implies that more extreme opinions can sometimes be correct, yet unpopular. Under such circumstances, credibility may erroneously be equated with popularity, and accuracy can be compromised when dissent is easily suppressed (see Lanier 2006 for a fuller critique of "the new online collectivism").

reliability. These forms of "citizen science" are credible by virtue of "credentialing," achieved through the training and education of novices by experts. Following specific training, citizens are invoked as sensors (Goodchild 2007) to provide accurate information about their local environments. As Goodchild has noted, the power of "humans themselves, each equipped with some working subset of the five senses and with the intelligence to compile and interpret what they sense" (p. 218) is a tremendous geographical resource. Key to the effective use of humans as sensors in specialized contexts, however, is establishing acceptable standards or metrics of credibility.

Research and tools informing the understanding of VGI credibility

Networked digital media simultaneously provide exciting opportunities for the provision of geographic information and formidable challenges to discerning its credibility. Yet, to date researchers have barely begun to examine the credibility of VGI. Pressing questions in this pursuit include whether users and professionals will accept systems populated largely by volunteered input as credible and, if so, for what purposes and with what effects? What factors impact users' credibility perceptions and thus technology use, acceptance, and information contribution? In what ways might geographic information differ from other types of content, such as health, commercial, or news information? And, what technical and sociotechnical tools can help users and professionals navigate VGI systems appropriately?

Research findings from other domains are a useful starting point in addressing these questions. For example, people tend to use cognitive heuristics—useful mental short cuts, rules-of-thumb, or guidelines that reduce cognitive load during information processing and decision making (Tversky and Kahneman 1974)—rather than more rational, but laborious, strategies to evaluate credibility (Hilligoss and Rieh, in press; Metzger et al. 2008; Sundar 2008). Heuristic cues such as the professionalism and layout of the information on web pages, ease of navigability, and absence of clear commercial intent (e.g., lack of advertisements; .gov or .edu domains versus .com) have been shown to positively impact credibility judgments (Metzger 2007), although complex interaction effects are present as well (Walther et al. 2004). Such findings, of course, have obvious implications for designers and users of new VGI and other GIS applications. Moreover, user reliance on a particular information resource is also positively correlated with credibility perceptions (Johnson and Kaye 2000), implying that as people become more familiar with Wikimapia or Google Earth, for example, they will likely perceive them to be more credible, irrespective of information regarding data quality (to a point, presumably).

In a study of how users determine the credibility of user-volunteered information (Metzger et al. 2008) users of Wikipedia were much less skeptical about its information quality than were nonusers. This study also found that people were somewhat sophisticated in their evaluations of the credibility of information gleaned from social computing applications. For example, people indicated they paid attention to the number of comments or reviews available in product ratings systems (such as in Amazon.com or epinions.com), as well as their valence as a set, and several expressed concern about the number of authors contributing to a Wikipedia entry. Most agreed that more testimonials, or more authors or contributors to a Wikipedia article, would produce less biased and, thus, more credible information. These findings suggest that as existing VGI systems increase in popularity and usage, both their perceived credibility, and perhaps even actual data quality, will increase over time. At the same time, however, problems of knowing what VGI systems and sources to trust will likely continue to affect usage of these systems.

Related social computing environments may also provide useful lessons for developing tools to help users evaluate VGI specifically. For example, several efforts are currently underway to assist users of Wikipedia in assessing the quality and trustworthiness of user-contributed information. WikiScanner, working on the assumption that credibility judgments properly rest on author identity, is a tool that reveals the identity (via IP address) of contributors and editors of Wikipedia content. A similar tool could be developed to increase author transparency or to indicate more fully the provenance of data sources within VGI systems. Indeed, Frew (2007) discusses the crucial role that author or source identification

plays in evaluating the quality of geographic information specifically.[4]

Other Wikipedia-based projects include edit history visualizations being developed at IBM and Xerox PARC, and Wiki Lab's Trust Coloring project (http://trust.cse.ucsc.edu/). Edit history visualization tools, such as Wikidashboard (http://wikidashboard.parc.com/), allow users to better understand not only who is editing Wikipedia entries, but how often those edits are made. This can reveal text longevity or editing patterns that may implicate credibility, for example when two individual contributors engage in a vigorous back-and-forth "edit war." Trust coloring provides Wikipedia users with credibility cues on a within-article, line-by-line basis. Textual trust is computed by considering the "reputation" of the author and text, as measured by longevity of his or her contributions in their original form (i.e., text by authors that is retained and is otherwise not changed or moved by others gains reputation). Text of varying reputation then appears in different colors within each Wikipedia page. Again, these projects could serve as prototypes for developing tools that are appropriate for VGI environments.

Implicit in the trust coloring system is the notion that leaving information unchanged is a form of content endorsement. Indeed, social endorsement is another important component of credibility (Flanagin and Metzger 2008; Fogg and Tseng 1999) and studies find that credibility judgments of online information are heavily influenced by others' evaluations of that information (Metzger et al. 2008). Applied to the VGI context, this argues for developing tools that incorporate others' assessments of the information—something akin to collecting and displaying users' "ratings" or approval of GIS information contributed by users (see, for example, Bishr and Kuhn 2007). Indeed, this is already apparent in Wikimapia, where the information contributions of new or unverified users must receive a positive vote to be retained. Social bookmarking, tagging, and markup tools operate under similar principles. These types of systems or tools attempt to more fully harness collective intelligence and social filtering that can be employed toward helping users discern credible geographic information.

[4] Most of the concerns about credibility in wiki environments, and especially Wikipedia, revolve around author anonymity, which may not be the primary issue with most VGI systems, many of which may not include author anonymity as a central feature. Still, other issues regarding credibility exist, even in non-anonymous authorship systems. For example, while authorship in VGI systems may be explicit and known to users, the methods used to collect and produce geographic information (e.g., the origins of the raw data, the programs run on those data to produce the maps, and so forth) may not be apparent and yet are critical to consider when evaluating the credibility of the data.

Research directions and conclusion

The proliferation of information sources via computer networks and the attendant rise of VGI suggest its potentially profound impact on the collective body of geospatial knowledge available today, and particularly in the future. Of critical importance in understanding this knowledge base are issues of information and source credibility and their relation to the production and use of VGI. Research endeavors addressing these issues can and should take many forms.

For example, there is a pressing need to study why people create and contribute to VGI repositories. The issue of contribution to "discretionary databases," where people can choose to contribute or to free-ride on others' information contributions, is a longstanding concern in the social sciences (see Marwell and Oliver 1993), and recent conceptualizations have articulated specific issues germane to the digital media environment (Bimber et al. 2005). With regard to VGI and credibility specifically, issues of contribution matter because people's motivations to contribute information have implications for its credibility. For example, people may contribute data to support others within a community of users, to achieve desired political outcomes, or because doing so is consistent with an altruistic world view of generalized reciprocity. Each of these motivations, though, suggests greater or lesser potential for bias or deception, and therefore has implications for information and source credibility.

In addition, research should take into account the specific nature of VGI. In the arena of user-generated political information (e.g., political blogs), for instance, people are aware of the subjective nature of the information contributed by users and likely factor this into their credibility assessments. However, some forms of VGI (such as Wikimapia, for

instance) ask laypeople to contribute information to what is traditionally thought of as "factual" or "objective" geographic data (i.e., maps). If people think of maps as providing objective information, do they then transfer their (positive) credibility assessments to Wikimapia, for instance? This suggests the wider question of whether people are *more* likely to believe VGI than other forms of user generated content because its products are often seen as providing more seemingly "objective" information. Similar notions of "credibility transfer" have been proposed across media outlets (Schweiger 2000), and seem appropriate to VGI as well.

Research is also needed on the changing and emerging market for VGI, in order to understand credibility and its implications in this context. While the market segments for traditional geographic information products and services (such as mapping agencies) have been well researched and, consequently, their requirements for reliable and accurate information are well understood, little is yet known about who uses VGI, for what purposes, and what their requirements for credible information are. For example, the question of how sensitive VGI users are to inaccuracies and misperceptions and, thus, how important the credibility of VGI and its suppliers is in this market should be addressed.

Measurement issues are also central to VGI credibility research, as noted by the contrasting views of credibility outlined here. For example, credibility can be assessed in terms of information "accuracy" (based on some accepted social standard or scientific reference point) for information that is primarily fact-based. Accordingly, research might address the best way to track the "provenance" of VGI, to ensure or at least make more explicit its origin and quality over time. This is particularly critical as information sources become murkier by virtue of more inclusive and often anonymous contributions, and as VGI gets reused and repurposed (e.g., through mashups and the like). However, for more opinion-based VGI, credibility should be assessed in terms of people's *perceptions* of information or source credibility. Indeed, Bishr and Kuhn (2007) go so far as to propose that because the quality of all geographic information is subjective to some extent, trust should be used as a proxy for VGI quality, and demonstrated in the form of user ratings. In their view, by making trust values explicit (via user trust rating systems), trustworthiness can be used as an alternative measure of information quality.

Finally, research needs to address the social, educational, and political outcomes of VGI by asking whether VGI assists, enables, or empowers citizens and citizen organizations (Miller 2006). For instance, use of GIS for electoral redistricting suggests the contentious and highly politicized use of geospatial data (see Eagles et al. 2000). As geographic information is increasingly user-generated, the potential for such highly politicized applications and manipulations is amplified. Specific research questions pertaining to the credibility of VGI might include whether VGI can empower individuals to be more geographically or environmentally aware and knowledgeable, or if VGI helps or hinders the establishment of social capital and community.

As the amount of VGI continues to grow, issues of credibility should assume a prominent place on the research agenda. Interdisciplinary collaboration drawing on fields including geography, information science, communication, psychology, sociology, and computer science are critical to understand the credibility of VGI, given the importance of collective knowledge about information processing and evaluation, technical considerations specific to geographic data, network capabilities and patterns, social influences, and user attitudes and motivations. Put another way, harnessing the collective intelligence of those concerned with networks of individuals sharing and evaluating geographic information—in essence, the wisdom of the academic crowd—may be the most powerful tool to understand how voluntarily contributed geographic information is and should be produced, evaluated, and used by specialists and novices alike.

References

Balram, S., & Dragicevic, S. (2006). Collaborative geographic information systems: Origins, boundaries, and structure. In S. Balram & S. Dragicevic (Eds.), *Collaborative geographic information systems* (pp. 1–23). Hershey, PA: IGI.

Bimber, B., Flanagin, A. J., & Stohl, C. (2005). Reconceptualizing collective action in the contemporary media environment. *Communication Theory, 15*, 365–388.

Bishr, M., & Kuhn, W. (2007). Geospatial information bottom-up: A matter of trust and semantics. In S. Fabrikant & M. Wachowicz (Eds.), *The European information society:*

Leading the way with geo-information (pp. 365–387). Berlin: Springer.

Bose, R., & Frew, J. (2005). Lineage retrieval for scientific data processing: A survey. *ACM Computing Surveys, 37*(1), 1–28.

Callister, T. A., Jr. (2000). Media literacy: On-ramp to the literacy of the 21st century or cul-de-sac on the information superhighway. *Advances in Reading/Language Research, 7*, 403–420.

Eagles, M., Katz, R. S., & Mark, D. (2000). Controversies in political redistricting GIS, geography, and society. *Political Geography, 19*(2), 135–139.

Elmes, G., Challig, H., Karigomba, W., McCusker, B., & Weiner, D. (2004). Local knowledge doesn't grow on trees: Community-integrated geographic information systems and rural community self-definition. In P. F. Fisher (Ed.), *Advances in spatial data handling* (pp. 29–40). Berlin: Springer Science and Business Media.

Elwood, S. (2006a). Beyond cooptation or resistance: Urban spatial politics, community organizations, and GIS-based spatial narratives. *Annals of the Association of American Geographers, 96*(2), 323–341.

Elwood, S. (2006b). Critical issues in participatory GIS: Deconstructions, reconstructions, and new research directions. *Transactions in GIS, 10*(5), 693–708.

Eysenbach, G. (2008). Credibility of health information and digital media: New perspectives and implications for youth. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility*. Cambridge, MA: MIT Press.

Eysenbach, G., & Kohler, C. (2002). How do consumers search for and appraise health information on the World Wide Web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal, 324*, 573–577.

Fischer, F. (2000). *Citizens, experts, and the environment: The politics of local knowledge*. Durham, NC: Duke University Press.

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism & Mass Communication Quarterly, 77*(3), 515–540.

Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of Web-based information. *New Media & Society, 9*(2), 319–342.

Flanagin, A. J., & Metzger, M. J. (2008). Digital media and youth: Unparalleled opportunity and unprecedented responsibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 5–27). Cambridge, MA: MIT Press.

Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems* (Vol. 1, pp. 80–87). New York: ACM

Frew, J. (2007). Provenance and volunteered geographic information. Retrieved March 10, 2008 from http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Frew_paper.pdf.

Fritch, J. W., & Cromwell, R. L. (2001). Evaluating Internet resources: Identity, affiliation, and cognitive authority in a networked world. *Journal of the American Society for Information Science and Technology, 52*(6), 499–507.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*, 211–221.

Google. (2007). Introducing Google Earth outreach. Retrieved March 11, 2008 from http://www.google.com/intl/en/press/pressrel/outreach_20070625.html.

Gunther, A. C. (1992). Biased press or biased public? Attitudes toward media coverage of social groups. *Public Opinion Quarterly, 56*, 147–167.

Harris Interactive. (2007). Four in five of all U.S. adults—an estimated 178 million—go online. Retrieved January 24, 2008, from http://www.harrisinteractive.com/harris_poll/index.asp?PID=827.

Hilligoss, B., & Rieh, S. Y. (in press). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing and Management*. doi:10.1016/j.ipm.2007.10.001.

Hovland, C. I., Janis, I. L., & Kelley, J. J. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.

Jackson, J. (2006). "Neogeography" blends blogs with online maps. *National Geographic News*. Retrieved March 6, 2008 from http://news.nationalgeographic.com/news/2006/04/0425_060425_map_blogs.html.

Johnson, T. J., & Kaye, B. K. (2000). Using is believing: The influence of reliance on the credibility of online political information among politically interested Internet users. *Journalism and Mass Communication Quarterly, 77*(4), 865–879.

Kyem, P. A. (2004). Of intractable conflicts and participatory GIS applications: The search for consensus amidst competing claims and institutional demands. *Annals of the Association of American Geographers, 94*(1), 37–57.

Lanier, J. (2006). Digital Maoism: The hazards of the new online collectivism. *Edge,* No. 183. Retrieved June 5, 2008 from http://www.edge.org/3rd_culture/lanier06/lanier06_index.html.

Lenhart, A. (2006). User-generated content. Retrieved January 24, 2008, from Pew Internet Project: http://www.pewinternet.org/PPF/r/76/presentation_display.asp.

Madden, M., & Fox, S. (2006). Riding the waves of "Web 2.0": More than a buzzword, but still not easily defined. Retrieved January 18, 2008, from http://www.pewinternet.org/pdfs/PIP_Web_2.0.pdf.

Marwell, G., & Oliver, P. (1993). *The critical mass in collective action: A micro-social theory*. New York: Cambridge University Press.

Metzger, M. J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology, 58*(13), 2078–2091.

Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. (2003). Credibility in the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In P. Kalbfleisch (Ed.), *Communication yearbook 27* (pp. 293–335). Mahwah, NJ: Lawrence Erlbaum.

Metzger, M. J., Flanagin, A. J., & Medders, R. (2008). Social and heuristic approaches to credibility evaluation. Manuscript submitted for publication.

Miller, C. C. (2006). A beast in the field: The Google maps mashup as GIS/2. *Cartographica, 41*(3), 187–199.

O'Reilly, T. (2005). What is web 2.0: Design patterns and business models for the next generation of software. Retrieved January 22, 2007 from: http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web–20.html.

Peluso, N. L. (1995). Whose woods are these? Counter-mapping forest territories in Kalimantan, Indonesia. *Antipode, 27*(4), 383–406.

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. In B. Cronin (Ed.), *Annual review of information science and technology* (Vol. 41, pp. 307–364). Medford, New Jersey: Information Today.

Schweiger, W. (2000). Media credibility—experience or image? A survey on the credibility of the World Wide Web in Germany in comparison to other media. *European Journal of Communication, 15*, 37–59.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the American Association of Geography, 96*(3), 491–507.

Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems, 32*(1), 1–5.

Sundar, S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. Metzger & A. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). Cambridge, MA: MIT Press.

Tseng, H., & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM, 42*(5), 39–44.

Tulloch, D. L. (2007a). Many, many maps: Empowerment and online participatory mapping. *First Monday, 12*(2). Retrieved March 7, 2008 from http://firstmonday.org/issues/issue12_2/tulloch/index.html.

Tulloch, D. L. (2007b). Position paper. Retrieved March 7, 2008 from http://www.ncgia.ucsb.edu/projects/vgi/docs/position/Tulloch_Paper.pdf.

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

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.

Walther, J. B., Wang, Z., & Loh, T. (2004). The effect of top-level domains and advertisements on health web site credibility. *Journal of Medical Internet Research, 6*(3):e24. <URL: http://www.jmir.org/2004/3/e24/>.

Weiner, D., & Harris, T. M. (2003). Community-integrated GIS for land reform in South Africa. *URISA Journal, 15*, 61–73.