Published in final edited form as:
Cartogr Geogr Inf Sci. 2014 ; 41(2): 115–124. doi:10.1080/15230406.2013.874200.

Spatial Collective Intelligence? credibility, accuracy, and Volunteered Geographic Information

Seth E. Spielman
Department of Geography, University of Colorado, 110 Guggenheim Hall/260 UCB, Boulder, CO, USA

Abstract

Collective intelligence is the idea that under the right circumstances collections of individuals are smarter than even the smartest individuals in the group (Suroweiki 2004), that is a group has an "intelligence" that is independent of the intelligence of its members. The ideology of collective intelligence undergirds much of the enthusiasm about the use of "volunteered" or crowdsourced geographic information. Literature from a variety of fields makes clear that not all groups possess collective intelligence, this paper identifies four pre-conditions for the emergence of collective intelligence and then examine the extent to which collectively generated mapping systems satisfy these conditions. However, the "intelligence" collectively generated maps is hard to assess because there are two difficult to reconcile perspectives on map quality- the credibility perspective and the accuracy perspective. Much of the current literature on user generated maps focuses on assessing the quality of individual contributions. However, because user generated maps are complex social systems and because the quality of a contribution is difficult to assess this strategy may not yield an "intelligent" end product. The existing literature on collective intelligence suggests that the structure of groups more important that the intelligence of group members. Applying this idea to user generated suggests that systems should be designed to foster conditions known to produce collective intelligence rather than privileging particular contributions/ contributors. The paper concludes with some design recommendations and by considering the implications of collectively generated maps for both expert knowledge and traditional state sponsored mapping programs.

Introduction

In the early twentieth century radio was unregulated. Individuals could buy small transmitters and broadcast messages from their home. These small transmitters only had a range of a mile or two but in the US alone over 2 million of them were sold by 1924 (Coll 2011). Nicolai Tesla thought the spread of these small transmitters and the resulting communication among distributed individuals meant that "the entire earth will be converted into a huge brain" (Wu 2010, p. 5). This idea, that allowing people to communicate with each other may lead to the emergence of some new form human intelligence, has re-emerged around in discussions of user generated content and the internet (O'Rielly 2005).

However, unlike radio where "intelligence" was an unanticipated externality, many Internet technologies are explicitly designed to harness "collective" intelligence (Madden and Fox 2006). Collective intelligence is the idea that under the right circumstances collections of individuals are smarter than even the smartest individuals in the group (Suroweiki 2004), that is a group has an "intelligence" that is independent of the intelligence of its members. The philosophy of collective intelligence is expressed in web based geospatial technologies through the design of systems that have low costs of entry. Easy entry into these systems is important because the philosophy of collective intelligence holds the quality of a product is inherently related to the number of contributors, more users and contributors leads to a better product (O'Reilly 2005; Hackalay 2010). Lanier (2010) notes that systems that espouse this philosophy explicitly privilege information quantity over information quality. The justification for this approach to information, according to Lanier, is that at some scale quantity begets quality.

However, this philosophy of collective intelligence is under-explored in the context of spatial information systems: What is spatial collective intelligence? Does collective intelligence exist in crowd sourced geographic information? Wooley (2010) has found that Collective Intelligence is not a universal property of collaborative groups, it emerges under specific circumstances. To what extent do online mapping systems satisfy these criteria? This paper reviews the current literature on the emergence of collective intelligence in groups and explores two questions: 1) What does collective intelligence mean within the context of the crowd sourced geographic information? 2) Do crowd sourced spatial information systems satisfy the conditions for the emergence of collective intelligence. The conclusion considers the short and long term implications of this philosophy for the "experts" that have traditionally generated geographic data (such as state sponsored statistical agencies).

What makes a good user generated map?

Web-based collaborative mapping systems, like Open Street Map, embody the collective intelligence philosophy. Open Street Map collects Volunteered Geographic Information (VGI) from a distributed set of users and assembles these contributions into a "patchwork" map (Goodchild 2007). To create the patchwork individual contributions of spatial data are aggregated (through some review process) into a single map. Goodchild and Li (2012) identify two review processes, one termed *social* and one called *crowd sourced* review – these two quality control processes are not mutually exclusive. In social review the quality of contributions are mediated or checked through a hierarchy of contributors. In *crowd sourced* review other users check contributions. Elwood et al. (2012) note the diversity of user generated web-mapping systems; some systems collect VGI and allow users to store information in personal maps that can be shared privately or with the public (such as Google). Still others record the locations from which messages are sent or check-ins and reviews of particular locations.

Researchers are often interested in considering these user generated content in the aggregate, that is the messages of specific individuals are not of interest, but the collected set of individuals is examined for geographic patterns or trends (Currid and Williams 2009). The process of creating a single patchwork by integrating inputs from volunteers raises questions about the reliability and quality of both the volunteered data and the aggregate cartographic end product; these questions form the basis of much of the literature on crowd sourced geographic information (Ellwood 2008). This discussion has, to date, centered around two different conceptions of quality. There are those who aim to formally evaluate the spatial accuracy of volunteered information (for example Hakalay 2010). These formal assessments are rooted in the concept of “spatial error” (Chrisman 1991) that considers the positional accuracy of geographic data, the accuracy of attributes, the completeness of the data, and other factors that with sufficient resources and time can be objectively assessed. Goodchild and Li (2012) suggest that the quality of individual contributions can be checked for accuracy against a set of rules that define plausible arrangements of geographic features – for example the physical propertied of a user contributed coastline can be checked to ensure that have a fractal structure as identified by Mandlebrot (1967). Mooney et al. (2010) take a similar approach to validation, arguing that new contributions should be geometrically similar to existing contributions of the same type, e.g. new lakes should have the same approximate geometric properties as existing lakes. Others argue that discrepancies in volunteered information may not constitute errors (Ellwood 2008, Elwood et al 2012). Different individuals can have different perspectives on the geographic world and the validity of these perspectives, and hence the validity of volunteered geographic information, can be judged based upon the credibility of the volunteer (Flanagan and Metzger 2008). In this perspective it is difficult to objectively assess the accuracy of the data but one can assess its credibility via the reputation, trustworthiness, and motivations of the contributor (Bishir and Mantelaas 2008).

Crowd-sourced data is the scientific equivalent of fast food, it is ubiquitous and cheap to acquire, but is it healthy and safe to consume (for research purposes)? The tension between the *validity-as-accuracy* and *validity-as-credibility* perspectives on crowd sourced geographic exposes one of the fundamental challenges for the use of these data for research and policy making purposes. Traditional statistical concepts of uncertainty and bias are very hard to apply and without these concepts the traditional models of accuracy disappear. These traditional tools are typically applied to see if data are “healthy” for consumers. Subjective measures of quality are also difficult to apply to crowd sourced data; in a large database it is difficult to examine the motivations and credibility of each contributor. Not all crowd-sourced information is equal, some data have higher quality than others (Hackalay 2010) and some contributors make better contributions than others.

Given that crowd sourced data are of varying quality, when aggregating such data one has to guard against regression to the mean. That is, highly accurate or highly credible contributions should be not be degraded by being combined with lower quality contributions, even if these exceptional contributions are outliers. A “good” user generated map would possess this quality and this is an essential characteristic of collective intelligence. However, the concept goes further than simply stating that good user contributions should not be “diluted” by bad contributions. Collective Intelligence implies that the end product is in some way better than the “best” individual contribution, because the group product enhances the individual contributions. If crowd sourced data are not better than the best individual contribution these systems are reducing human expertise, not enhancing it.

Without collective intelligence maps produced by experts should be better than crowd sourced maps, with collective intelligence experts are obsolete. There is another lighter definition of collective intelligence which sets a lower bar for crowd sourced data, which is that of the product of the group is independent of the individual contributions of data. Another way of stating this rule is that the whole map must be something independent from the sum of its parts, that the group as a whole has created something more than a collection of individual contributions. This definition raises some interesting and underexplored challenges, such as, the notion that filtering individual contributions for quality may not yield the "optimal" collectively generated end product. Notwithstanding debates about the criteria for quality, a better understanding of collective intelligence and how it interacts with individual inspiration, expertise, and creativity can only enhance the use of aggregated crowd sourced data.

Defining collective intelligence

O'Rielly (2005) notes, "an essential part of Web 2.0 is harnessing collective intelligence, turning the web into a kind of global brain…" The full O'Reilly definition is quite long but in general suggests that the web 2.0 is "an attitude not a technology." This attitude focus on user interaction, "Web 2.0 companies ***set inclusive defaults for aggregating user data and building value as a side-effect of ordinary use of the application*** (emphasis in the original) … they build systems that get better the more people use them." This user-generated data is the "Intel inside" web 2.0 systems, as a result these systems require developers to "trust" their users (O'Rielly 2005). This is a nice, democratic, design philosophy, but it is predicated on specific ideas about the net effect of large-scale human interaction, that is interaction between people produces "intelligence." More specifically, the interaction of individuals around some specific task produces a super-individual (group-level) intellect.

Until recently there was little more than anecdotal evidence in support of the existence of group (collective) intelligence. In *Science* Wooley (2010) reported the identification of a "Collective Intelligence Factor" (Wooley 2010). The Collective Intelligence Factor is a property of the group itself, not merely the sum of the intelligence within the group. Wooley et al. (2010) through a series of experiments that involved examining the performance of groups on specific tasks (like playing checkers against a computer and solving visual puzzles) found that the intelligence of groups is largely independent from the intelligence of their members. That is, having intelligent people in the group was not a significant predictor of group performance.

While it is possible to assess group performance on well defined, experimentally controlled tasks, applying this work to crowd sourced spatial information raises the difficult question of the quality of maps—what is an intelligent map -- one could possibly attain very different measures of quality by focusing on the credibility or the accuracy of the end product. The simplest form of assessment is positional accuracy, and by that criterion there is evidence of collective intelligence, as the number of contributions increases map accuracy increases (Hakalay 2010). Perhaps collectively generated on-line maps, because of their broad user base, are inherently credible (even if they are less spatially accurate than the most accurate user contribution)?

Without a specific portrait of map quality spatial collective intelligence is difficult to define. This does not mean that the question of collective intelligence in crowd sourced spatial data cannot be meaningfully considered. If individual contributions to mapping systems are often superior to the collectively generated end product than these systems degrade the accuracy of information. On the other hand if the collectively generated end product is better than the "best" individual contribution these new information systems are incredibly valuable. It may be that by the credibility criterion online maps necessarily generate collective intelligence because their credibility improves as the number of contributors increases. Spatial Collective Intelligence is a largely subjective term; its key premise is that collections of individual contributions are by some metric "better" than existing spatial data resources.

The concepts of credibility and accuracy are different from the concept of collective intelligence in that they apply to both individual user generated content and the resulting collectively generated product. By contrast collective intelligence, as a metric can only be applied to the end product, not individual contributions. An "intelligent" user generated map is one that is either credible, accurate, or both. Credibility and accuracy should not be conflated; collectively generated maps may produce a distinct type of collective intelligence-one that is credible but not necessarily accurate.

One could argue that user generated maps are inherently credible because of the size of their user base. A map can be credible without reflecting the real world, for example Crutcher and Zook (2009) show that post-Katrina, online mapping systems reproduces digital divides in the broader society, African-American neighborhoods in New Orleans received less online attention than other neighborhoods. Grinberg et al. (2013) using Twitter and Foursquare data found that in the days following Hurricane Sandy in New York City the number of people going out to bars and clubs to party increased. This seems like an odd activity for those whose homes were recently destroyed. Crawford (2013) notes that few of the tweets and check-ins studied by Grinberg et al. came from the areas hardest hit by Sandy, most came for affluent areas of Manhattan where the young, wealthy, and well connected reside. In these instances, the aggregate picture fails to uniformly represent the reality of lived experience- in spite of these systems large user bases.

From the existing literature it is clear that not all groups are intelligent. Understanding the nature of collective intelligence and the ability of crowd sourced spatial information systems to generate it is key to assessing the potential societal impact of crowd sourced spatial data and its fitness for various forms of use. Such assessment requires a clear enumeration of the conditions that create collective intelligence.

Criteria for Collective Intelligence: Aggregation and Diversity

Collective intelligence emerges only under certain circumstances; it is not an automatic property of groups. Some groups, because of their structure and communication style, are more intelligent than others (Wooley 2010). Wooley et al. (2010) found that the most important factors in collective intelligence are the social sensitivity of group members and egalitarian communications within the group (measured as the variance in the number of speaking turns by group members). Soroweiki (2004) argues that for collective intelligence to emerge 4 criteria are necessary:

- **i.** Independent contributors
- **ii.** Diversity of opinion
- **iii.** Decentralization
- **iv.** A way to aggregate the results.

Wooley's (2010) research seems to support Suroweiki's (2004) criteria. Of these criteria only the third, decentralization is satisfied on its face the by the collective GeoWeb. The independence of contributors, the aggregation of results and the diversity of volunteers on the websites that collect volunteered data warrant careful attention.

Diversity of opinion

It is difficult to assess the diversity of online communities because their users are essentially anonymous. Elwood et al. (2012) note that little is known about why people contribute geographic information. At the time of writing we are aware of only one systematic survey describing the characteristics and motivations of VGI contributors. Budhathoki and Haythornthwaite (2013) sent a survey to all of the 31,105 active contributors to Open Street Map via the Open Street Map messaging system. They collected 444 valid responses. Responders were overwhelming young (64.6% between 20 and 40), male (96%), and college educated (95%). The Majority of Budhathoki and Haythornthwaite (2013) respondents contributed to open-source software projects (60.3%) and/or Wikipedia (71.5%).

Open Street Map is in many ways substantively similar to Open Source Software projects that that ask users to voluntarily contribute time and expertise without expectation of direct material compensation. Indeed one of the most striking features of **Budhathoki** and Haythornthwaite (2013) survey is the demographic parallels between Open Street Map and Open Source Software projects- both have a the massive gender imbalance in contributors. A non-random sample of contributors to the Linux kernel found that 96% of the 141 survey respondents were male (Hertel at al. 2003). A larger study of the Open Source Software community a drew a random sample of projects that were under active development and conducted a census of finalized products on SourceForge a repository for collaborative programming projects resulting in a sample of 684 contributors, a 30% response rate, providing data on 287 different Open Source projects (Lakhani and Wolf 2003). Lakhani and Wolf (2003) found that respondents were overwhelming male (97.5%) with a mean age of 30 (SD 8 years, the youngest contributor was 14 the oldest was 56). The gender imbalance in these systems is troubling because in an experimental setting Wooley (2010) found that the number of females in a group was a significant predictor of its intelligence, Wooley attributes this to women's generally higher level of sensitivity to others.

However, the diversity criterion does not principally apply to demographics, the opinions and motivations of contributors are essential. There is a substantial body of research on the motivations of contributors to open source software. Generally, the research finds that individual motivations are diverse, some are motivated to contribute by the love of programming, others enjoy the sense of community, and others are motivated by political beliefs. The contributor's top 3 motivations were diverse, some wanted to improve programming skills (41.3% selected as a top 3 motivation), others found programming intellectually stimulating (44.9%), but the most common motivation was needing the collectively generated code for a separate commercial or non-commercial project (58.7%). Bitzer et al (2007) argue that one of the most important factors leading to the creation of an Open Source Project is perceived need, projects are started to address gaps or shortcomings in existing software systems. The same seems to be true for Open Street Map, the most successful projects often address a pressing need or areas of poor geographic coverage such as crisis situation (e.g. Haiti post-earthquake).

The motivations to contribute to Open Street Map are similar. Budhathoki and Haythornthwaite (2013) through a factor analysis of survey responses found that a latent variable measuring the expectation of direct or indirect monetary rewards from contribution was the most important factor for contribution among respondents, this factor accounted for 20% of the variance in the sample. However, other factors identified by Budhathoki and Haythornthwaite (2013) relate to needing the data for commercial or non-commercial project – and additional 10% of the variance was explained by factors relation to personal need for the data and the desire to promote oneself. The second most important factor was "learning" and this too parallels Lakhani and Wolf's (2003) findings. The third most important factor identified by Budhathoki and Haythornthwaite (2013) was "self-efficacy regarding local knowledge" accounted for 8% of the variance in their sample and measured both the sense of enjoyment derived from contributing and the ability to improve data about a place for which the respondent had local knowledge. Variables related to a person's beliefs were identified, belief in the goals of the Open Street Map project and a sense of altruism explained an additional 12% of the variance.

Wikipedia also asks users to volunteer information to a collectively generated product. Yang and Lai (2010) studied 200 English language contributors to Wikipedia and found that "internal self-concept motivation" was the most important predictors of the frequency of individual contributions. Yang and Lai define internal self-concept motivation as a form of positive feedback that occurs when a behavior is consistent with an individual's ideals and provides positive external feedback. People contribute to collective knowledge bases when they find the activity consistent with their ideals ***and*** they receive some praise for doing so from others. Yang and Lai's (2010) analysis seems to concur with Budhathoki and Haythornthwaite (2013) findings.

The research shows that the motivations for contributing to Wikipedia and Open Source Software, and Open Street Map are diverse but the users who make contributions are not (they're overwhelming young men). People contribute to address existing gaps in systems and because doing so can provide tangible financial benefits, reinforce political beliefs, provide a sense of community, and/or positive reinforcement of their beliefs. Systems that collect VGI have characteristics of both Open Source Software and Wikipedia. The data suggests that individual motivations for participating may be diverse even if the individuals are of a similar age and gender.

Independent contributors

A diverse group is a necessary but not sufficient condition for the emergence of collective intelligence; diversity is capitalized upon through egalitarian communication (Wooley, 2010). In this regard web based mapping systems are problematic, participation in web based mapping systems is best described by a long tailed distribution in which a small number of contributors account for a large portion of contributions, a few contributors dominate the conversation (Sieber 2010, Haklay et al 2008, Ellwood et al 2012, Goodchild and Li 2012). The communication imbalances that seem to characterize web 2.0 mapping systems potentially stifle the ability of these systems to generate collective intelligence. Wooley's (2010) work shows that when the loudest person in the group dominates the conversation intelligence suffers.

Members of an intelligent group exhibit independence and diversity. Historically speaking, large gatherings of people around a specific purpose often exhibit the opposite sort of behavior - uniformity and interdependence. A physical (or virtual) crowd is a purposeful group, a special type of crowd that is guided by a specific cause, event, or purpose. In *Crowds and Power* (1960), a book that spans history, philosophy, and literature, the Nobel Laureate Elias Canetti identifies four attributes of crowds:

1. Crowds always want to grow
2. Within a crowd there is equality
3. The crowd loves density
4. The crowd needs a direction

Some of these attributes, like density, do not apply to virtual crowds but others, such as the need for direction and growth seem relevant to the web 2.0. Canetti is deeply skeptical about the nature of crowds; however, this skepticism should be offset by the now documented existence of collective intelligence. In Canetti's taxonomy of crowds, websites like Open Street Map could be characterized as a "slow crowd." A slow crowd is different from other types of crowds: Flight crowds are fleeing en masse from a threat; Baiting crowd are out to harm a specific person (or group), in both of these cases the crowd's goal (safety or violence, respectively) is attainable. Slow crowds on the other hand are motivated by a remote, immaterial and unattainable goal. As a result slow crowds are denied "discharge", the satisfaction of reaching their goal. The GeoWeb is a slow crowd because The Map of a dynamic world is never complete.

While the goal of online crowds may be remote, they are not without direction. Direction in online crowds emerges in interesting ways, Flanagan and Metzger (2008) note that, "credibility judgments of online information are heavily influenced by others' evaluation of that information" (p 145). This kind of following behavior can be seen as an example of online crowds' direction seeking behavior; it is the antithesis of independence. Direction, the movement of crowds toward a single goal hinders diversity and independence. The acceptance of mass opinion as a measure of credibility emphasizes aspects of crowd behavior that stand in direct opposition to the characteristics of "smart" crowds. When, as Flanagan and Metzger (2008) describe, credibility is influenced by others assessment of credibility- credibility grows through the direction seeking behavior of crowds not via the merit of the contribution relative to other (newer) contributions; when following behavior exists contributions are not independent. The weight of prior assessments drives the credibility of contributions. This is a particular challenge for VGI, in order for collective intelligence to emerge these systems need to find a way to foster egalitarian communications and counteract the following tendency of crowds. Systems that privilege frequent contributors, or those that have contributed credible information in the past, by assigning their contributions a high base level of quality, run the risk of inducing following behavior. One might call this a credibility loop—the credibility of a contribution is self-reinforcing.

A way to aggregate the results

For group intelligence to emerge one needs a way of combining individual contributions into a collective end product. Financial markets are a particularly efficient means for this, Suroweiki (2004) describes the response of the stock market to the loss of the Space Shuttle Challenger as an example of collective intelligence. Prior to an investigation of the shuttle disaster the NY stock market devalued the stock price of the company that produced the part that failed – according to Suroweiki (2004) the market's "intelligence" identified the offending part manufacturer.

In most existing examples of collective intelligence the method of aggregation is self-evident; this is not the case with maps. In financial markets all bids are equal and asset prices are the aggregate of bidding behavior. However, when it comes to VGI all contributions may not be equal and this makes the task of aggregating results into a single map quite difficult. In a financial market all bids are equal because those unable to credibly bid are excluded from the market via the high costs of entry. However, the web 2.0 in general and sites that collect VGI in particular, by design limit the cost of entry to facilitate broad-based participation.

There is evidence that barriers to participation may be an important part of intelligent groups (Galton 1907). One of the most prominent pieces of evidence in support of collective intelligence is Sir Francis Galton's 1907 article in Nature, *Vox Populi*. In this article Galton studied entries from a carnival game that asked contestants to guess the weight of a dressed ox carcass. Galton found that while individual tickets tended to misstate the weight of the ox considered as a whole the group's guess at the ox's weight was remarkably close to its true weight. The group's median guess of 1207 pounds was astonishingly close to the true weight of 1198 pounds (a difference of less than 1%).

Viewing Galton's experiment as evidence in support of the wisdom of all crowds is a mistake. The cost of the contest was integral to success of the crowd, Galton explicitly notes that the cost of entering the contest (six pennies) was high enough to deter "practical joking" (p. 450). Even though contributions of group members were simple integers written on a ticket Galton was deeply concerned with how to aggregate these tickets into a collective group estimate. For Galton the exclusion of jokers was important because measures of central tendency, like the mean, are sensitive to outliers. Galton's concern about jokers can also be understood as a concern about the credibility of entrants, a joker is an entrant without credibility. Galton's crowd seems to satisfy the previously discussed aspects of collective intelligence: Contestants were diverse, they included butchers and farmers as well as people with little experience with livestock. However, in spite of their diversity they were all were (independently) pursuing the same goal (the prize). Unlike Galton's experiment, individuals' contribution to web 2.0 mapping systems may not be guided by a common purpose and the costs of entry are deliberately minimized (O'Reilly 2005).

Goodchild and Li (2012) discuss the strategies currently used to aggregate features and/or reconcile multiple versions of the same feature. When mediated collectively generated mapping systems use one or more of three strategies which they term *social, crowd sourced, or geographic* (Goodchild and Li 2012). Open Street Map, for example, a crowd-based follows a last-in, first-out model. The features on the collectively generated map that users of the site see contain the location and attributes of the last person to edit that feature. In Open Street Map aggregation functions like a stack, the most recent edit is one users see. This form of aggregation does not directly leverage prior contributions and relies heavily on trust, however Hackalay (2010) found that more edits to an Open Street Map feature generally leads to more positional accuracy. Social strategies rely on a hierarchy of users. If conflicts arise Open Street Map uses a *Social* mechanism in which a set of elite users adjudicate problems. Finally, Geographic strategies seek to validate a feature based upon its geographic context and or geometric properties- for example, an intersection can only occur at the junction of one or more streets.

Cartographic products are positioned and political, maps are incapable of including all information or telling all possible stories (Monmonier 1996; Wood 1992). The political nature of maps may be manifest online in any number of ways, for example, feature selection. The evidence suggests individuals contribute information to online systems when they stand to derive some sort of benefit and/or when they feel there is a gap in existing systems and contributing information is consistent with their ideals. The combination of these motivations makes mapping seem like an intensely political exercise.

The political nature of mapping and the aggregation of user inputs potentially lead to an oddly inconsistent mash-up of worldviews. For example, Bishop (2011) argues that since the 1970s in the United States at the national, regional, and local scales place is increasingly associated with sets of political beliefs. If feature selection is political, and political outlooks are geographically clustered, the aggregation of crowd-sourced data is problematic. Monomier (1995) discussed the mosaicking of data with different ages in a GIS as resulting in system where one moved through time as one moved through the space depicted on a map. Moving through a patchwork of assumptions and perspectives seems deeply problematic, especially when these perspectives and assumptions are invisible to end-users.

Aggregation is a critical aspect of collectively generated map products and its implications have to be carefully considered. Without effective means of aggregation maps will either be shaped by the most active contributors or map features will reflect simply reflect the average of contributions. Most aggregation strategies focus on evaluation individual contributions of data (Goodchild and Li, 2012). It is not clear that strategies that focus on individual contributions necessarily lead to the best collective product. If collectively generated maps are viewed a complex systems in which individual behavior is shaped by the behavior of others its possible that the best collective product is not attained by selecting the best individual contribution. The optimal result could be contingent upon (or incorporate) low quality contributions.

User generated maps do not have to undergo aggregation. Prior to large state sponsored geographic and social surveys hobbyists in the US would write gazetteers (Cohen 1981). These gazetteers were regional geographic encyclopedias. Individuals would systematically document the history, population, economic activity, and location of towns in a state or county. Sometimes these documents have mind-numbing levels detail, recounting the number of cows in a township. Gazetteer's were created at significant personal expense and took years to compile, often in the course of creating the gazetteers authors would visit every town in a region, conduct interviews with town elders, and gather data through correspondence with businesses (Cohen 1981). In some sense web 2.0 mapping systems are new versions of these gazetteers however, there is one critical difference. The gazetteers of the 18th century had single authors who signed their works web 2.0 systems often combine data from multiple authors raising very difficult questions about validity, quality, and aggregation.

Do online maps meet these criteria?

By the diversity criteria the ability of web 2.0 systems to produce intelligent map products seems mixed. On the one hand contributors may have diverse motivations for volunteering, on the other hand contributors may lack demographic diversity. These systems have exceptionally low costs of entry in order to ensure broad based participation but in spite of the low costs the evidence suggests that the contributor base may not be broad- the majority of contributions come from a minority of contributors. The downside of these low barriers to entry is that the credibility of contributions is difficult to assess. Imposing higher costs, such as training and/or affirmation of codes of conduct might reduce the number of contributors but would increase the credibility and/or accuracy of contributions.

These systems do not possess the egalitarian communication structure of effective groups. Using the Goodchild and Li (2012) schema both *social*, and *crowd sourced* data validation strategies rely on authority or the order in which contributions are received. One of the downsides of the reliance on the authority of users in crowd-based systems is that it can be self reinforcing (Flanagan and Metzger 2008), this type of feedback plays into the properties of crowds identified by Canetti (1962). To some extent these shortcomings can be addressed through design, these systems should aim to recruit sustained communication from a diverse user base and efforts can be made to encourage egalitarian communications as opposed to privileging the opinions of frequent contributors.

Some of the critical questions about user-generated maps arise because it is difficult to devise systems that aggregate individual contributions of geographic data into a single collectively generated end product. Is aggregation and quality assessment a technical problem or fundamental problem with the ideology of crowd wisdom that underlies web 2.0 geospatial technologies? The answer to this question may lie in the concept of spatial collective intelligence. While this concept is difficult to define precisely careful attention to the criteria outlined above may yield cartographic products that capture the promise of the web 2.0 and ensure that collaboratively generated maps do not degrade human expertise or exceptional insights into the geographic world.

The implications of collaborative user generated maps

In the web 2.0 geospatial data infrastructure volunteers serve as both “citizen sensors” and “citizen censors”. Individuals are both the creators and arbiters of information. This is a significant shift from a time when state sponsored mapping and survey organizations were the only game in town. While state sponsored mapping programs are by no means free from individual perspectives and biases (Brodaric 2004), a shared goal, norms of behavior, and a desire to maintain the credibility of the institution allows the work of many individuals to be seamlessly combined into credible and accurate cartographic products. Shearmur (2010) notes the concepts and techniques used in large-scale public sector data collection efforts “have been arrived at though a slow cultural consensus, and because the sampling framework, survey methods, questions, and definitions are all in the public domain (and thus open for debate) that the census has authority unlike other statistics.” While the web 2.0 software systems are open to varying degrees, they are not products of slow cultural consensus. These systems have rapidly “matured” and in spite of their flaws they are on the verge of being “locked-in” (Lanier 2010).

Flyvberg (1998) notes, “Power determines what counts as knowledge, what kind of interpretation attains authority as the dominant interpretation.” In the web 2.0 power is expressed through software systems that structure individual interaction. Collectively generated maps express ideas about the nature of the world. The ambiguities of the real world tend to get reduced in software systems and these reductions get “locked-in” (Lanier 2010). Lock-in is the process where a software system’s view of the world becomes reality. That is the representation of the world embodied in the software system becomes so widely accepted that it is difficult to imagine alternatives. In Geographic Information Science an example of lock-in is the concept of layers. Thinking of the world as a collection of layers of points, lines, or polygons is a conceptual trope, it is one of many possible ways to represent geographic features. Layers are a way of reducing the complexity of the real world, it is entirely possible to conceptualize geographic features as a mix of points lines and polygons or as fields of objects (Cova and Goodchild 2002) however this form of representation is “locked-out.” After nearly 30 years of use the concept of layers is locked-in, it is how we represent the geographic world in computers. Lanier (2010), drawing from Karl Popper, notes that the scientific method disqualifies thoughts as it proceeds; the disqualification is based on a deliberative process of peer-review. In contrast to scientific processes lock-in disqualifies ideas based on what is easy to program, politically feasible, fashionable, or created by chance. Lock-in makes philosophy real. In spite of their flaws locked in ideas tend to persist; the cost and effort associated with the design of software systems is a formidable source of inertia. Due to this inertia locked-in technologies are buffered against the process of peer-review (Lanier 2010). Consider the minimal impact of the decades of literature on temporal GIS and temporal data models on locked-in GIS software systems. The question of collective intelligence in online mapping systems should be subject to scientific scrutiny before crowd-based GIS get locked-in; Hackalay (2010) has made important progress in this area but more work needs to be done.

The development of technology that allows individuals to contribute and share spatial information about the places to which they are attached is a positive development. However, these developments can become problematic when they are aggrandized as way to conduct citizen science and “measure the whole world” (Goodchild 2007). In a recent editorial in the *Urban Geography* Richard Shearmur (2010) warned of, “the role that academic fashions may inadvertently play” in paving the way for political decisions such as the abolition of the state sponsored mandatory social surveys. In the 1970’s the academic discipline of Urban Planning “had theorized about itself to such an extent it that it was denying it’s own claim to legitimacy” (Hall 1996, p 334). Cartographers and Geographic Information Scientists should be wary of doing the same- crowd-based online systems with low costs of entry may devalue their expertise. If quality maps can be created and maintained on the fast and cheap from crowds in the clouds, why should the resource-constrained public sector duplicate the crowds’ effort?

Digital mash-ups of administrative information are commonly proposed alternatives for large-scale state sponsored surveys. For example in 2001 the Netherlands did a “virtual census” without the use of enumerators (Nordholt 2004). The USGS is currently considering incorporating VGI into the national map (Wolf et al. 2010). Movements in this direction should be watched carefully. If these new forms of data do not exhibit collective intelligence they represent a loss of information, a loss of expertise, and a devaluation individual creative and scientific inspiration. Credible and accurate geographic data, like those produced by national statistical agencies are critical for civil discourse and should be defended. While crowd sourced data may be credible, simply because of the size of its user base, one has to ask careful questions about both its spatial accuracy and its potentially complex biases.

Conclusions

Web 2.0 geospatial systems represent a new and exciting form of human communication about the environment. Currently, much of the literature on VGI is focused on the assessment of the quality of individual contributions, this focus is appropriate. However, the critical question is how to ensure that the collection of contributions is high quality; this may or may not be contingent on the quality of individual contributions- more research is necessary. It is entirely possible that the way to maximize the quality of the collectively generated end product is to focus on designing systems that aim to generate (or support) group intelligence. Such systems would focus on diversity, egalitarian communication structures, and mechanisms to ensure that the aggregate leverage prior contributions. Systems are currently dominated by young men and designed to minimize the cost of entry. Perhaps development efforts should go into voting mechanisms that adjudicate claims collectively instead of via the social mechanisms described by Goodchild and Li (2012). It may also be possible expose users motivation for contributing by automatically building profiles that describe the places and the kinds of features they edit, such automatically generated profiles would expose motivations and allow users to better understand the authors of particular features.

Designing better collaborative mapping systems simply by focusing on the quality of contributions is difficult to do because the quality of a map, or a contribution to a map, is difficult to determine. The validity-as-accuracy and validity-as-credibility perspectives are both reasonable criteria and are difficult reconcile. Ensuring that systems are well designed may be a better way to ensure the quality of the cartographic end product than filtering/evaluating individual contributions. The web 2.0 design philosophy holds that with appropriate processing, massive quantities of information from a group of anonymous contributors can be translated in very high quality information (Lanier 2010, O'Rielly 2005). The problem of applying this philosophy to the geographic realm is that it is entirely unclear how the collection and "processing" should proceed. The literature on collective intelligence suggests that it matters how this information is collected and how it is processed. By focusing on designing systems that explicitly foster collective intelligence to promise of user generated maps may be more effectively realized.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Bishop, B.; Cushing, R. The big sort: Why the clustering of like-minded America is tearing us apart. Mariner Books; 2011.

Bitzer J, Schrettl W, Schroder P. Intrinsic motivation in open source software development. Journal of Comparative Economics. 2007; 35(1):160–169.

Budhathoki NR, Haythornthwaite C. Motivation for Open Collaboration Crowd and Community Models and the Case of OpenStreetMap. American Behavioral Scientist. 2013; 57(5):548–575.

Canetti, E. Crowds and Power. New York: Farrar, Strauss, and Giroux; 1962.

Chrisman N. The error component in spatial data. Geographical information systems. 1991; 1:165–174.

Coll S. The Internet: For Better or Worse. The New York Review of Books. 2011 Apr 7. 2011.

Cohen P. Statistics and the State: Changing Social Thought and the Emergence of a Quantitative Mentality in America, 1790 to 1820. The William and Mary Quarterly. 1981; 38(1):35–55.

Cova T, Goodchild M. Extending geographical representation to include fields of spatial objects. International Journal of Geographical Information Science. 2002; 16(6):509–532.

Crutcher M, Zook M. Placemarks and waterlines: Racialized cyberscapes in post-Katrina Google Earth. Geoforum 40.4. 2009:523–534.

Elwood S. Volunteered geographic information: future research directions motivated by critical, participatory, and feminist GIS. GeoJournal. 2008; 72(3):173–183.

Elwood S, Goodchild MF, Sui DZ. Researching volunteered geographic information: Spatial data, geographic research, and new social practice. Annals of the Association of American Geographers. 2012; 102(3):571–590.

Flanagin A, Metzger M. The credibility of volunteered geographic information. GeoJournal. 2008; 72(3):137–148.

Flyvbjerg, B. Rationality and power: Democracy in practice. University of Chicago Press; 1998.

Galton F. Vox populi. Nature. 1907; 75:450–451.

Goodchild M. Commentary: whither VGI? GeoJournal. 2008; 72(3):239–244.

Goodchild M. Citizens as sensors: the world of volunteered geography. GeoJournal. 2007; 69(4):211–221.

Goodchild MF, Li L. Assuring the quality of volunteered geographic information. Spatial statistics. 2012; 1:110–120.

Grinberg, N.; Naaman, M.; Shaw, B.; Lotan, G. Extracting Diurnal Patterns of Real World Activity from Social Media; Proceedings of the International Conference on Weblogs and Social Media 2013; 2013.

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

Haklay M, Basiouka S, Antoniou V, Ather A. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus Law to Volunteered Geographic Information. Cartographic Journal. 2010; 47(4):315–322.

Hall, P. Cities of Tomorrow. Oxford: Blackwell; 1996.

Hertel G, Niedner S, Herrmann S. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. Research policy. 2003; 32(7):1159–1177.

Lakhani K, Wolf R. Why hackers do what they do: Understanding motivation and effort in free/open source software projects. MIT Sloan Working Papers. 2003

Lanier, J. You are not a gadget: A manifesto. Vintage; 2011.

Levy D, Peart S. The Tale of Galton's Mean: The Influence of Experts. Public Choice. 2002; 113(3): 357–365.

Madden M, Fox S. Riding the waves of Web 2.0. Pew Internet & American Life Project. 2006:1–6.

Monmonier, M. How to lie with maps. University Of Chicago Press; 1996.

Monmonier, M. Drawing the line: Tales of maps and cartocontroversy. Henry Holt & Co; 1995.

Mooney, P.; Corcoran, P.; Winstanley, AC. Towards quality metrics for openstreetmap; Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems; 2010.

Nordholt, E. The Dutch Virtual Census 2001: A New Approachby Combining Different Sources. UNECE Seminar on New Methods for Population Censuses; 2004.

O'Rielly T. What Is Web 2.0. O'Rielly Media. 2005 http://oreilly.com/web2/archive/what-is-web-20.html.

Sieber R, Rahemtulla H. Model of Public Participation on the Geoweb. GIScience Proceedings 2010. 2010

Shearmur R. A World Without Data? The Unintended Consequences of Fashion in Geography. Urban Geography. 2010; 31:1009–1017.

Suroweicki, J. The Wisdom of Crowds. Anchor; 2004.

Wood, D.; Fels, J. The power of maps. The Guilford Press; 1992.

Woolley A, Chabris C, Pentland A, Hashmi N, Malone T. Evidence for a collective intelligence factor in the performance of human groups. Science. 2010; 330(6004):686. [PubMed: 20929725]

Wu, T. The master switch: the rise and fall of information empires. Knopf; 2010.