Deconstructing the binaries of spatial data production: Towards hybridity
Jonathan Cinnamon

Alt.GIS special issue of *The Canadian Geographer*

*Binaries, the most reductive form of categorization, can be usefully invoked to characterize emerging phenomena; yet, they are widely critiqued for oversimplifying a complex world and for their use as tools of social and political influence. Through a literature review and content analysis this article traces the emergence of volunteered geographic information (VGI), and identifies the recurrent use of several related binaries to contrast this phenomenon with the conventional spatial data production activities of states and corporations. Using several key examples, these binaries are deconstructed by identifying a mismatch in how VGI is conceptualized (bottom-up, amateur, asserted) in the literature and the reality of existing VGI projects. As an alternative to a binary conceptualization of spatial data production, a different representation is put forward that more accurately depicts what is in actuality a vast, shifting, and heterogeneous landscape of spatial data production approaches. Thinking about contemporary spatial data production not as a binary but as a continuum could encourage the development of hybridities that harness the benefits of different approaches—including the oversight and quality control of conventional methods, with the speed, low cost, and distributed nature of citizen-based spatial data production.*

Keywords: Spatial data, binaries, volunteered geographic information, critical GIS, crowdsourcing
Introduction

Once in place, difference and boundaries between the new and old perspectives have a strong tendency to rigidify… (Kwan 2004, 757)

The 1990s were a time of polarizing debate over the place of geographic information systems (GIS) in geography. On one side, proponents of GIS touted its role in legitimizing geography as a ‘scientific’ discipline (e.g., Openshaw 1991), while on the other side, numerous critics—mainly human geographers but also some members of the GIS community—identified a host of concerns, including its circumscription of methodology, ontology, and epistemology at a time when the discipline was evolving and pluralizing in these areas (e.g., Taylor 1990; Smith 1992; Pickles 1995). Further, GIS was criticized as an agent of marginalization due to its high cost and complexity, which served to restrict its use by some groups (e.g., Harris and Weiner 1998; Obermeyer 1998). These “GIS and Society” debates grew to become a core research theme of geographic information science (GIScience), which itself evolved to become critical GIS (Schuurman 1999), a field of inquiry aimed at enlightening the GIS and geography communities to the technology’s socio-political contestations and democratic challenges. Also as part of critical GIS, a number of GIS and mapping practices were developed to address these challenges. Public participation GIS (PPGIS) (Corbett and Keller 2005; Sieber 2006) has flourished as an alternative to conventional GIS practices that are sometimes anti-democratic, technocratic, and exclusionary (Obermeyer 1998). Similarly, ‘counter-mapping’ projects (Peluso 1995; Harris and Hazen 2006)—which seek to represent space in order to highlight the interests of the socially or politically disenfranchised—were positioned in diametric contrast to dominant hegemonic mapping by the state and other structures of authority.

While initially common and helpful to illustrate the need for a critical understanding of maps and GIS, the once widely employed binaries of critical GIS—e.g., empowerment/marginalization, GIS/GIS-2, mapping/counter-mapping—are now less frequently invoked. The longer-term success of critical GIS has been shaped not by the continuation of binaries but by hybridity and plurality, especially due to the advancement of novel forms of thinking in critical GIS. Drawing from emerging ideas on situated knowledges, subjectivity, reflexivity, and positionality, and from a wider disciplinary interest in alternative and non-representation, mixed-methods, and ontological and epistemological hybridity, the recently-emerging critical GIS sub-areas of feminist GIS and qualitative GIS are emblematic of the shift from binaries to more pluralist understandings in critical GIS (Knigge and Cope 2006; Kwan and Knigge 2006; Elwood 2008a Cope and Elwood 2009). This maturation, I would argue, has been responsible for the sustained success of critical GIS, once the dust from the two-sided debate had settled. This argument has also been put forth by Nyerges et al. (2011), who in the introductory chapter to their edited book on GIS and Society, argue that:

[The debate about the ‘good’ and ‘evil’ sides of geographic information technologies continues to this day, though in a much more qualified form. The naivety of sharply polarized thinking begins to unravel quite readily once one begins more fully to research, deconstruct, and reconstruct the]
relationships at play. Simplistic perspectives on either side of the debate are being superseded by an understanding that the intents and actions of agents need to be situated within their social and institutional contexts in order for GIS technology developments and use to be judged as good or evil. This more mature stance is part of a growing interest in creating an intellectual foundation for approaching the development and use of GIS as being itself a complex societal process (5).

Analogous to the initial use of binaries in GIS and Society and then critical GIS, the geospatial Web (or GeoWeb) (Scharl and Tochtermann 2007) has been similarly positioned in contrast to desktop GIS, as an alternative, if not overly antagonistic geospatial domain (Crampton 2010). The GeoWeb broadly refers to the collection of mapping platforms and geospatial products and services available on the Web, which are typically available in accessible packages designed for broad, public usage. Akin to the strategy of GIS and Society and critical GIS, binary conceptualizations have been useful to describe the GeoWeb in relation to desktop GIS; however I contend that this dualistic image is no longer necessary nor is it accurate. This has been particularly evident for depictions of volunteered geographic information (VGI) (Goodchild 2007a)—a social phenomenon closely tied to the GeoWeb—which has been positioned as the binary opposition to conventional spatial data production activities by states and corporations.

In this article I argue that spatial data production has been conceptualized using binary oppositions since the emergence of VGI in 2007 as an area of research within GIScience. Although constructive and essential at the outset to characterize a nascent phenomenon, their persistence may become problematic. Given the significant and sustained interest in VGI, I contend that it is now time to move beyond these binaries, which could promote productive, pluralist engagements and heterogeneous spatial data production activities. The trajectory of critical GIS as sketched out above could provide a template for the development of a future spatial data production defined not by oppositions, but by diversity and plurality. The article is outlined as follows: in the next section I briefly discuss the uses and the problems of binaries and categorization, followed by a literature review and content analysis of the VGI literature which illustrates the recurrent use of three related binaries (authoritative/asserted, expert/amateur, top-down/bottom-up) to conceptualize spatial data production in the age of VGI. The main section of the article then deconstructs these binaries by identifying a mismatch in how VGI has been conceptualized in the literature, and the reality of several high-profile VGI projects. In doing so, I demonstrate the inadequacy of these binaries for framing contemporary spatial data production activities. Next, I propose a different conceptualization that could better reflect the diversity of spatial data production. This section then emphasizes the hybridities that could be enabled by the cross-fertilization of conventional spatial data production with citizen- or volunteer-based approaches. In the conclusion section, I summarize the article’s argument and discuss some of the likely challenges that hybrid approaches to spatial data production may pose.
The binaries of spatial data production

The social construction of categories

The making of categories has been investigated in diverse fields, from neuroscience to science, technology, and society (STS) studies. From a neuroscience perspective, the use of binaries—the most reductive form of categorization—has been described as a basic artefact of human thought and behaviour rooted deep in our unconscious. From this perspective, the brain is hardwired to react to stimuli in a binary way—as ‘good’ or ‘bad’—which initiates a pre-cognitive behavioural response and quick decision-making, and is thus intimately tied to human survival (LeDoux 1994; Wood and Petriglieri 2005). Although categories allow us to survive in and make sense of the world, they can also be destructive. The benefits of reduction and simplification enabled by categorization are counterpointed by a world and its contents limited to a finite number of possibilities. Categorization and especially binaries limit the possibilities for what can be conceived about the world—as Woodward et al. (2009) describe, “binaries presume a totalizing epistemology, so termed because either/or thought can only posit a world in which everything either ‘is’ or ‘is not’.” (399). Alternatively, when informed by STS, categories and binaries are recognized as social rather than neurological constructs, which calls attention to their agency, materiality, and politics. Categories, like all socio-technical artefacts, emerge and crystallize due to myriad negotiations amongst various socio-technical actors, yet they are produced and concretized by specific actors to protect their own interests (Callon 1991; Akrich and Latour 1992). As Bowker and Star’s (2000) work has demonstrated, they are not neutral; rather, categories and systems of classification are themselves “powerful technologies” (147) that perform political work, often invisibly. Accordingly, it is important to recognize the materiality of categories and classification systems in order to understand their complex historical and political contexts (Bowker and Star 1996).

STS has proven to be a useful theoretical lens in the domain of spatial data and GIS (see Harvey and Chrisman 1998; Schuurman 1999; Harvey 2001; Wilson 2009), including for understanding the agency and materiality of objects such as categories or metaphors (Poore and Chrisman 2006). As Poore (2003, 66) explains, “[i]n a large socio-technical system such as GIS, systems, software, people, organizations, regulations and even inanimate objects are seen as playing active roles”. In the case of spatial data production, binary categories—given their materiality and agency—will exert influence on the entire socio-technical system in which they exist. The use of binaries to conceptualize the contemporary spatial data production landscape could alter, for example, how mapping technologies are designed, how data practices unfold and are operationalized, the composition of relevant policies and institutions, and how various actors in the system interact, or—which is important to the argument put forth in this paper—how they influence actors not to interact. Certainly, it can be said that binaries may be helpful as “first approximations” (Warf 2007, 210); however they must be identified and challenged, both due to their agency, materiality, and political power, and due to their tendency to limit ontological and epistemological diversity. Below, the
binaries used to conceptualize VGI are exposed via a review and content analysis of the VGI literature.

Reviewing the VGI literature: 2007-2013
While geographic information has always been produced in the public realm, interest in VGI—as an area of inquiry within GIScience—was triggered by a 2007 article written by Michael Goodchild (2007a) in which the author coined the term and in doing so initiated the process of conceptualizing the phenomenon of geographic information produced by citizens. Despite the widespread uptake and pervasive use of the term VGI, it is important to acknowledge that the ‘volunteered’ in VGI has been met with criticism since much of this information is collected and used without the knowledge of the person(s) who produced it (Elwood et al. 2012). An example of this is the scraping of georeferenced social media content ‘harvested’ by researchers for use in mapping projects, typically without the knowledge of potentially million of contributors (Stefanidis et al. 2013). For a good overview of that debate, see the chapter by Harvey (2013) in the edited book on VGI by Sui et al. (2013). Indeed, that debate remains active and has some resonance with the present argument, however the review of the VGI literature in the following paragraph does not attempt to judge the appropriateness of the use of the term ‘volunteered’. Regardless, VGI has since quickly become a major research topic, evidenced by a number of edited collections, conference workshops, and specialist meetings devoted exclusively to it. Whether truly volunteered or not, the term VGI has been used to describe a range of different types of geographic content produced by citizens, from GPS traces collected and uploaded to OpenStreetMap to geotagged photographs on Flickr or georeferenced tweets on Twitter, to citizen-contributed location information and points-of-interest on WikiMapia. VGI is often cited as a consequence of a broader increase in public engagement and of shifting information economies that came about due to the transformation of the Web from an information source to a platform that can enable democratic inclusion, interaction, and participation—a shift often referred to as ‘Web 2.0’ (O’Reilly 2005). In conjunction with the wider Web 2.0 shift, the advent of GeoWeb platforms designed for the average Web user (e.g., Google Earth) and mobile computing have both facilitated the rapid growth of VGI in the past several years.

A complete review of all the published and unpublished VGI literature was beyond the scope of this article; as such, the sampling strategy focused on collections of English-language VGI literature. Further, a number of standalone works were included. This targeted sampling strategy focused on overviews, reviews, and theoretical and conceptual pieces primarily, as it was assumed that they would be more likely to provide a conceptualization of VGI compared to pieces that mostly describe VGI applications or case studies. The list of works consulted include the 30 position papers from an agenda-setting VGI workshop held in Santa Barbara, California in 2007; 10 articles from a 2008 special issue of GeoJournal on VGI; 9 articles from a 2010 special issue of Geomatica on VGI; and 20 chapters of a 2013 edited book on VGI. Moreover, a number of standalone works on VGI during this time were included in the content analysis, including two foundational articles by Goodchild, and 29 other pieces from this period, collected via a literature search for ‘volunteered geographic information’ as the keyword. The author
hand-searched all of the texts, looking for descriptions, definitions, or conceptualizations of VGI.

The review identified that the VGI literature has continually engaged with a series of related binaries to conceptualize this phenomenon in contrast to conventional spatial data production, based around the either/or binary structure. These binaries are top-down/bottom-up, expert/amateur, and authoritative/asserted. The persistent use of these binaries from 2007 to 2013 is illustrated below using a number of examples from the reviewed VGI literature.

Goodchild’s influential introduction to VGI (2007a) set off a wave of interest in the phenomenon, and in the same year the author provided a particularly unambiguous binary conceptualization of VGI (2007b):

The worlds of VGI and the traditional mapping agencies could not be more different. The latter represent the top-down, authoritarian, centrist paradigm that has existed for centuries, in which professional experts produce, dissemination is radial, and amateurs consume. Expertise in this world is measured with objective indicators such as advanced degrees; progress requires consensus and is therefore slow and deliberate; and costs rise steadily. The world of VGI is chaotic, with little in the way of formal structures. Information is constantly being created and cross-referenced, and flows in all directions, since producers and consumers are no longer distinguishable (29).

Goodchild’s dualistic description likely had an influence on future commentators. Following Goodchild’s introduction, a VGI workshop was held in Santa Barbara at the end of 2007. The three identified binaries were used in some of the participants’ position papers to differentiate VGI from conventional spatial data production and spatial data infrastructures (SDI). In outlining their position on VGI, authors invoked the binaries of top-down/bottom-up (Gould 2007), expert/amateur (Craglia 2007), and authoritative/asserted (Goodchild 2007c) to situate VGI in opposition to conventional spatial data production. An influential 2008 special issue of GeoJournal helped to define the research agenda for VGI in subsequent years. In this issue’s commentary piece, Goodchild suggests that VGI must be understood in contrast to data production by national mapping agencies, and that doing so “explains why VGI is so interesting, provokes such a strong reaction in many quarters, and stimulates the kinds of writing represented in this special issue” (2008, 239). In that special issue, Flanagan and Metzger (2008) framed their exploration of the credibility of VGI using an expert/amateur binary, Elwood (2008a) described how nascent research in this area used the bottom-up/top-down pairing to differentiate VGI from SDIs, while Bishr and Mantelas (2008, 236) claim that “[VGI] has emerged from the bottom up and does not tend to rely upon top-down approaches to ensuring information quality or data sharing.” Also in 2008, Sui (2008) employed top-down/bottom-up and expert/amateur binaries when describing VGI and related phenomena, to illustrate the challenges and implications of the “wikification of GIS”.

The use of language that positions VGI on one side of an either/or opposition to conventional spatial data production continued in subsequent years, including
contributions to a 2010 special issue of Geomatica on VGI. An article in this issue by Grira et al. (2010) puts forth a framework for understanding VGI in the context of spatial data infrastructures. This framework is based on a binary separation of what the authors call “authoritarian” and volunteered spatial data, in which authoritarian refers to authoritative data produced by experts, and volunteered refers to asserted data produced by amateurs. Also in 2010, Genovese and Roche (2010) utilize the top-down/bottom-up binary to contrast VGI with SDIs, while Warf and Sui (2010) describe a “profound shift in the nature and role of geographic information, a transition characterized by a ‘bottom up’ reconfiguration in how data are collected, transmitted, analyzed, visualized, and utilized that differs considerably from traditional ‘top-down’ models in which experts and government agencies dictate the criteria of data collection…” (197).

Although VGI is now a well-established field of inquiry, more recent works continue to use the same binary conceptualizations. In 2012, Elwood et al. (2012, 574) state that VGI “can be said to be asserted, in contrast to the authoritative products of traditional sources that derive their authority from their creation by highly trained experts” [emphasis in original], and Leszczynski (2012, 75) argues that “[c]ontrary to a model where GI is produced by cartographic experts and disseminated by them to passive end-users, VGI efforts represent a distinctly different regime, whereby it is private citizens—increasingly addressed as consumers—who create spatial data over the Web”.

In 2013, an edited book on VGI was published, entitled “Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice.” (Sui et al. 2013a). Based on the contributions to this volume published six years after Goodchild first described the phenomenon, binary conceptualizations are still used. The book’s overview chapter argues that “in contrast to the traditional top-down authoritative process of geographic data production by government agencies, citizens have played an increasingly important role in producing geographic data of all kinds through a bottom-up crowdsourcing process” (Sui et al. 2013b, 3). The introduction to Jiang’s chapter (2013, 125-26) in this book states that “[t]he emergence of VGI represents something of a paradigm shift in terms of geographic data acquisition from the conventional top-down approach, mainly dominated by national mapping agencies, to the bottom-up approach, in which data are contributed by individual volunteers through crowdsourcing…” Additionally, several authors in this volume make use of the authoritative/asserted and expert/amateur binaries (Feick and Roche 2013; Johnson and Sieber 2013). Also in 2013, Fairbairn and Al-Bakri (2013, 350-51) describe a “distinct contrast between those who capture and collate geospatial data as members of the public … and those state-sponsored or commercial organizations which provide official or Formal Data.”

Conceptualizing VGI in binary opposition to conventional spatial data production has been necessary for describing the phenomenon. As Elwood notes with respect to naming debates over VGI itself, they “are important in part because they focus our attention on those concepts that are key to understanding this new phenomenon, and help us identify critical issues for this emerging research agenda” (2008b, 133). Although binaries are useful to encourage interest in a nascent phenomenon, the main contention of this article is that their continued use may cause more harm than good, since once fully embedded, binaries restrict possibilities for productive cross-engagements. The following section deconstructs these binaries by illustrating how they are inadequate for framing
what is in actuality a heterogeneous range of contemporary spatial data production activities.

**Deconstructing spatial data production’s binaries**

Although binaries have routinely been invoked between 2007 and 2013 to position VGI in opposition to conventional spatial data production, this conceptualization often does not match the reality of many contemporary spatial data production activities. In this section I deconstruct these binaries by demonstrating the inappropriateness of terms such as ‘bottom-up,’ ‘amateur,’ and ‘asserted’ to conceptualize some spatial data production activities widely described as VGI. Conversely, I also illustrate how some conventional spatial data production activities by states and corporations might not fit the narrative of authoritative data produced by experts.

Although Goodchild’s influential paper stated that VGI “is chaotic, with little in the way of formal structures” (2007c, 29), this description is inaccurate for several of the highest profile examples. OpenStreetMap (OSM), certainly one of the best-known and celebrated VGI projects since it began in 2004, has been portrayed in binary opposition to top-down corporate products such as Google Maps and those of the national mapping agencies such as Natural Resources Canada, Ordnance Survey in the United Kingdom, and the US Geological Survey (USGS). Yet, in certain ways OSM resembles them quite closely. While anyone is free to contribute and edit data, spatial data production via OSM is typically quite structured (Feick and Roche 2013) and resembles authoritative initiatives in many ways (Eckert 2010). For instance, OSM data are frequently collected using the ‘mapping party’ concept. Using this model, knowledgeable organizers, including OSM insiders, typically recruit members of the public, provide them with instruction on how to collect street data, and often carry out the data handling, uploading, and editing tasks (Perkins and Dodge 2008; OpenStreetMap 2012a). Oversight and quality control are core aspects of the mapping party model—much as they are to conventional spatial data production—despite the portrayal of VGI as bottom-up, informal, and ‘chaotic.’

A closer look also reveals the top-down corporate influence on OSM. CloudMade, a company founded by Steve Coast who also founded OSM, develops application programming interfaces (APIs) and mapping products that exploit citizen-created OSM data (OpenStreetMap 2012b). CloudMade at one point also had employees known as ‘Community Ambassadors’ who raised awareness about OSM and organized mapping parties, all in the name of monetizing the OSM dataset (Eckert 2010). The once-ubiquitous MapQuest, owned by AOL, is involved with OSM via MapQuest Open an OSM-based alternative to their main Web mapping site. MapQuest Open can also be accessed within the main OSM Website, presenting, with the click of the mouse, an ostensibly more cartographically-appealing rendering of the OSM data than that of the standard version. As Leszczynski (2012) describes, fuelling corporate entities such as CloudMade or MapQuest through free citizen labour is an example of ‘wikinomics,’ a business model centred on profiting from crowdsourced Web content. Although this approach is still emerging with regard to citizen-generated map data, it is being increasingly leveraged in the world of ‘Big Data’ more broadly (see Thatcher 2014). This
Commercial developers of OSM now have their own dedicated meeting, the OSM Professional Large Users Summit (PLUS), organized by Steve Coast and other OSM insiders. The general focus of the meeting is to explore opportunities for enhancing OSM for commercial purposes, which includes considering if there is “a way forward to fund OSM with earmarked donations to achieve specific goals” (OSM PLUS 2013). This top-down/bottom-up merger of citizens and corporations might diminish the appeal of OSM and move it closer to closed, corporate alternatives such as Google Maps, yet, as Perkins and Dodge (2008) identified, this may be a necessary step to ensure the ongoing success of the project. Indeed, the success of OSM might be attributed to its blending of bottom-up and top-down strategies (see Eckert 2010).

Two closely related variations of the top-down/bottom-up binary also remain persistent in descriptions of VGI; expert/amateur, referring to the assumed level of expertise of the data producer, and authoritative/asserted, pertaining to the provenance and therefore perceived credibility of the data. A more nuanced inspection, however, reveals that VGI is produced by people with a range of backgrounds and proficiencies under a variety of circumstances. In reality, the credibility of data created using citizen-based approaches varies greatly, with some examples in particular possessing the characteristics of ‘authoritative’ datasets produced by ‘experts.’ Many citizen science and crowdsourcing projects have a volunteer-based spatial data collection component, ranging from large scale and distributed examples such as the Christmas Bird Count and the GLOBE program, to smaller one-off local data collection projects. GLOBE is a US-based, international environmental science and education program for school-aged students. Although the spatial data collected by students have frequently been held up as an archetypal example of VGI (e.g., Fritz et al. 2009; Sabone 2009; Elwood et al. 2012), GLOBE’s scientist-designed protocols are highly formalized and structured and extensive training and support materials are provided to students and teachers. The high level of training and expert involvement means that the data collected are considered authoritative enough to be used in scientific research and publications (GLOBE 2009). GeoWiki engages citizens around the world to help classify land cover maps using their own local knowledge and Google Earth imagery. Although this project is described in the context of VGI, Fritz et al. (2012) identified that the vast majority of the project’s current 66,000 contributions came from scientists, frequently with a background in remote sensing or GIS, even when widely advertised beyond these communities. The project, though relying on the ‘wisdom of the crowds,’ is perhaps better described as an “expert-sourcing system than a crowdsourcing (sic) one” (Fritz et al. 2012, 122).

Many well-known crowdsourcing VGI projects rely on the contributions of experts, including spatial data production efforts initiated in the aftermath of natural disasters or crisis events by groups such as Crisis Mappers and the Standby Task Force (SBTF) (see Li and Goodchild 2010; Roche et al. 2013). The SBTF, an international association of currently 900 volunteers from over 70 countries, links up with localized humanitarian efforts to provide highly structured geospatial support (Standby Task Force n.d.). SBTF volunteers are experienced crisis mappers who are oriented and trained in the specific needs of each deployment; this includes collating reports of the locations of disaster damage or political violence and adding them to Web-based maps, or annotating recent high-resolution satellite imagery with the locations of military equipment in conflict zones. The organizers and the network of volunteers interact through social
media and communication technologies, which fosters rapid uptake and allows spatial data and maps to be produced quickly and efficiently. As a typical example of the work that these volunteers do, the SBTF coordinated a global network of 250 volunteers to collect and map information pertaining to the political crisis and social uprising in Libya in 2011, using the well-known Ushahidi crisis-mapping platform (Standby Task Force 2011). SBTF volunteers from around the world collated instances of humanitarian issues from media reports and online social networks, and added them to the crisis map (see Figure 1). Despite being produced by volunteers, the spatial data and maps produced by SBTF and other ‘expert-sourcing’ initiatives ought to be considered authoritative since they are used for logistical planning and decision-making by the organizations that requested their assistance, including the United Nations Office for the Coordination of Humanitarian Affairs, Amnesty International, and the World Health Organization (Standby Task Force 2012). These examples of VGI produced in crowdsourcing and citizen science projects suggest that describing spatial data produced using these approaches as ‘asserted’ and the volunteers as ‘amateur’ serves to oversimplify a complex and evolving spatial data production activity.

![Ushahidi crisis map for Libya](image)

**Figure 1: Ushahidi crisis map for Libya.** Spatial data and maps were produced by volunteer members of the Standby Task Force (SBTF), a global network of crisis mapping experts. These were then used in official decision support by the agency partnered with the SBTF, the UN Office for the Coordination of Humanitarian Affairs (UNOCHA).

In addition to concealing the potentially significant expertise of volunteers through describing these data as asserted and produced by amateurs, these binaries also unjustly
presuppose the converse—that conventional spatial data producers are unquestionably authoritative. As Coleman (2013) contends, describing conventional spatial data production by governments as authoritative is not always based in reality: “[d]espite claims to the contrary, government topographic mapping products in most countries are not ‘authoritative’ by any practical definition. Far from it. They are typically out of date, possibly inconsistent, and usually the victims of diminishing maintenance budgets” (259-60). As suggested by Coleman’s assertion, we should not automatically ascribe authority to state and corporate spatial data producers. Goodchild (2009) has described the example of misregistered spatial data and imagery in two ostensibly authoritative corporate and state cartographic products, imagery in Google Earth and the US Geological Survey’s National Map. Another higher profile case comes courtesy of the world’s (currently) most valuable company, Apple. The 2012 release and immediate catastrophe of Apple Maps introduced millions of iPhone and iPad users to a core area of GIScience—spatial data accuracy and uncertainty—that in this case had “potentially life-threatening” consequences for drivers relying on it for navigation (see Thompson 2012). Although Apple is not a traditional producer of cartographic products, their financial clout and expertise in the world of technology ascribes authority to its products, despite its venture into the world of spatial data proving shockingly amateur. Indeed, even the design of the Apple Maps logo and app icon generates questions about their authority; the image in the icon suggests an impossible navigation direction—turning left onto Interstate 280 from a position on top of an overpass crossing over the freeway (see Figure 2).

Maps
Maps take a whole new turn.

Figure 2: The Apple Maps icon. A screenshot of the Apple Maps icon and catchphrase “maps take a whole new turn,” from the Apple iOS6 Web site (http://www.apple.com/ios/maps/). Note that the icon is suggesting an impossible navigation option, turning left directly from the overpass onto Interstate 280 below, bringing into question the authority of this product while also making the catchphrase strangely fitting (the link above now directs to the iOS7 website, which uses a newly styled and cartographically-appropriate logo).
Figure 3 shows the Lions Gate Bridge in Vancouver as represented in Apple Maps 3D mode. The obvious problem is that the imagery is quite significantly distorted (a common problem in Apple Maps); a less obvious problem is that the road dataset is incomplete, the bridge is labelled only as Highway 1A when it should also properly be designated as part of Highway 99. Interestingly and conversely, in Google Maps the bridge is incompletely identified solely as part of Highway 99. In OSM however, the bridge is correctly labelled as a segment of both Highway 99 and Highway 1A (see Figure 4). Although this is just one example of so-called ‘asserted’ data being more accurate than ‘authoritative’ data, there is some evidence pointing to the increasing convergence of citizen-generated data with official geographic data in terms of coverage and data quality. Although highly variable depending on location, the quality and coverage of OSM data has been shown to be comparable to those of some national mapping agencies (see Girres and Touya 2010; Haklay 2010). In the rapidly growing and diversifying world of spatial data production, these examples point to the hazards in unquestioningly assuming authority, based solely on the data producer’s perceived level of expertise, whether they are citizens, states, or corporations.
Figure 3: Vancouver’s Lions Gate Bridge in Apple Maps
Figure 4: Lions Gate Bridge in Google Maps and OpenStreetMap. The left image is a screenshot of the Lions Gate Bridge in Google Maps. Note that it is labelled as Highway 99 only. The right image is a screenshot of the bridge in OpenStreetMap, which correctly labels the bridge as a segment of both Highways 99 and 1A.

Hybridity: from binaries to synergies

Whatever we do we should not roll over and go back to sleep. Only the binary dreams of modernism and Western philosophy await us, and not the hybrid world of our waking lives (Barnes 2005, 76).

Given the examples described above, perhaps reconceptualizing contemporary spatial data production as a continuum rather than an either/or binary may be more reflective of reality. A continuum encourages greater ontological diversity compared to binaries, and potentially, heterogeneity and cross-fertilization in spatial data production. Figure 5 illustrates spatial data production as a continuum. On the extreme left, amateur citizens with no expertise in the subject matter or the use of geospatial technologies produce spatial data voluntarily in a fully bottom-up fashion; on the extreme right, expert employees of national mapping agencies produce spatial data in a fully top-down, authoritative fashion. Generally, cartographic upstarts such as Google and Apple should likely occupy a position on the continuum close to the national mapping agencies given the importance of spatial data quality and accuracy to the success of their cartographic operations; however, this is certainly up for debate given the examples provided above. While the continuum does allow for discrete binary oppositions, its value lies in the ability to represent a vast range of possibilities between the polarities. In the middle of
the continuum examples informed equally by each polarity are located, fully hybrid approaches. Leftwards from the middle are examples of approaches that might be described as increasingly bottom-up, amateur, and asserted. Moving rightward from the middle are examples of approaches that might increasingly be described as top-down, expert, and authoritative. Drawing on the ‘map-use cube’ (MacEachren 1994; MacEachren and Kraak 1997), a ‘spatial data production cube’ with three axes derived from the continua illustrated in this paper (authoritative—asserted, top-down—bottom-up, expert—amateur), might be a useful visualization tool for those designing hybrid spatial data production activities, or for comparing different spatial data production examples. Also, a continuum might be useful in the debates surrounding the suitability of the term ‘volunteered’ in VGI, by using Harvey’s (2013) separate definitions of VGI and ‘contributed geographic information’ (CGI) as the two polarities of a solely citizen-generated spatial data continuum.

Figure 5: Spatial data production as a continuum

Hybridity has emerged as a core principle for challenging the sanctity of existing categories and promoting diversity and cross-fertilization. Hybridity has been conceptualized in different ways; for instance, it has been described as “those things and processes that transgress or disconcert binary terms that draw distinctions between like and unlike categories of object—such as self/other, culture/nature, animal/machine or mind/body” (Whatmore 2009, 351). Within geography, a number of commentators have urged for a reconsideration of normative conceptual understandings that rely on rigid binaries, via a turn to hybridity. Kwan’s (2004) commentary identified two related binaries bifurcating the discipline—the separation of the human and physical sides of the discipline, and more specifically the divide between spatial-analytical and social-cultural geographies. Taking a pragmatic tone, the author suggests that “hybridity may be a productive stance for negotiating difference among geographers and geographies…the fluid identities it allows can be beneficial for creative geographical research” (Kwan 2004, 756). Recent articles by Sui and DeLyser (Sui and DeLyser 2012; DeLyser and Sui 2014) interrogate the qualitative-quantitative divide in geography. The authors highlight boundary crossings (such as the use of social theory in physical geography) and mixed method approaches (such as qualitative GIS) that call into question this enduring binary.
In describing these emerging disciplinary trajectories, the authors call for synthesis and holism to engage qualitative and quantitative polarities. Slightly differently, Sarah Whatmore (2002)—in her book on the nature/society binary—sees hybrid geographies as about more than the synthesis of two polarities. The author’s relational ontology of hybridity highlights the spaces in between the poles, all the while assuming partiality, mutability, and contingency (see Demeritt 2005). A relational approach to hybridity then, proposes more than the merging of binaries—it allows for true heterogeneity and the development of entirely new ontological possibilities, an exciting proposition for spatial data production.

Although spatial data production has been described in binary terms since the emergence of VGI, hybridity is a growing strategy in the public and private sectors and in academia. In the public sector, national mapping agencies are looking at the spaces between citizen-generated data and their own official spatial data production activities. The USGS has, since the 1990s, explored the potential value in citizen augmentation of the National Map through the Map Corps program. The USGS has organized the OpenStreetMap Collaborative Program (OSMCP), a phased, carefully planned path towards hybrid spatial data production (Wolf et al. 2011). The intention is to have citizens contribute to the National Map; however, given the importance of accuracy for this product, the USGS is carefully designing protocols for data production, quality control, and volunteer training to make certain the volunteered data are suitable for integrating with existing data (Wolf et al. 2011). Seeing as a main goal of SDIs is to coordinate the creation and use of spatial data between spatial data producers and users (Coleman and Nebert 1998; Masser 2005; Budhathoki et al. 2008), insights from this area may provide a foundation for facilitating these hybridities in the public sector, as will work from related areas of GIScience more broadly. For example, advances in data and metadata standards, semantic interoperability, and formal ontologies for integrating heterogeneous spatial data (e.g., Frank 2001; Fonseca et al. 2002; Schuurman and Leszczynski 2006; Buccella et al. 2009) have generated technical knowledge that is germane to the present challenge. Engagements with STS in this domain have illustrated how these technical challenges are inextricably bound up with the social and the institutional (Harvey and Chrisman 1998; Martin 2000; Chrisman 2005). The challenges of data integration and sharing between governments and diverse groups in society has been an important area of research (e.g., Poore 2003; Harvey and Tulloch 2006; Elwood 2008c), findings of which could inform the development of socio-technical solutions for hybrid spatial data production. Harvey’s (2003) work highlighting the importance of trust in the successful implementation of SDIs—given the variety of human and non-human actors involved—is pertinent, since trust is clearly also of fundamental importance in the case of spatial data production by heterogeneous contributors.

Utilizing these insights on spatial data sharing and integration forged over the past several decades will prove fruitful for enacting the emerging hybrid data production landscape. For example, there is a growing literature applying the knowledge gained from research on SDIs to the emerging challenge of data integration for VGI and crowdsourced data (e.g., Craglia 2007; Budhathoki et al. 2008; Cooper et al. 2010; Genovese and Roche 2010). Du et al. (2012) developed a formal ontology for the integration of road data from conventional and citizen-generated data sources, while Ramos and Devillers (2013) are exploring options for dealing with the semantic
heterogeneity issues posed by data produced by diverse contributors. Also, Palmer and Kraushaar (2013) have used STS and actor-network theory to understand the production of storm-report VGI by heterogeneous actors in a socio-technical sensor network, and its integration and use by the US National Weather Service.

Meanwhile in the private sector, companies are employing a wikinomics strategy by augmenting their spatial datasets with data produced by their users (Feick and Roche 2013). TomTom’s MapShare allows users of their navigation devices to provide updates to the spatial database, such as new roads, speed limits, or roadway construction (TomTom 2013). All users of TomTom products can then decide on what updates they consider to be authoritative—contributions by individual users, contributions provided by multiple users, or updates that have been verified by TomTom. Users of Google Maps can ‘Report a problem’ on the Website or the mobile app. Moreover, Google’s MapMaker program proactively engages citizens to update their cartographic products, although contributions must be authenticated before they are added to Google Maps and Earth (Google 2012). Apple Maps currently also has a ‘Report a Problem’ feature on their mobile app, but no large scale user-based updating scheme such as those of Google or TomTom.

Like the examples of hybridity in the public and private sectors, the spaces in-between citizen and official spatial data production could also prove to be fertile ground for the academic and scientific realms. Early excitement about the potential for vast amounts of data produced by globally distributed citizen sensors was quickly tempered by concerns over data quality, which threatened to relegate VGI to a mere curiosity rather than a viable source of scientific data. Recently hybrid solutions are emerging that try to combine the best of both worlds; the rapid speed and local knowledge of data produced by the world’s distributed citizens with the oversight, quality control, and data standards of conventionally-produced spatial data. Connors et al. (2012) call for hybridity to enhance environmental monitoring, and describe a platform designed to allow both average citizens and research scientists to contribute data pertaining to the spatial spread of a disease in trees. Also, Seeger (2008) described ‘facilitated-VGI,’ a hybrid approach that could be useful for producing scientifically valuable spatial data based on the knowledge of citizens. With this approach, volunteers are sometimes recruited—often based on their knowledge of the particular phenomenon of interest—and then asked to interact with mapping interfaces either face-to-face with facilitators, or remotely, to provide information according to predefined questions or criteria (Seeger 2008). Facilitated VGI projects rely on volunteer citizen-based knowledge, yet a carefully planned data production strategy helps to ensure the validity and credibility of the data for use by researchers and scientists. This approach has been utilized in research projects both with distributed volunteers (e.g., Fritz et al. 2009), and in face-to-face interactions (e.g., Thompson et al. 2011; Cinnamon and Schuurman 2013).

**Conclusion**

Although categories are useful for managing complexity, their very existence serves to limit the possibilities for understanding the world. Moreover, understanding categories as social constructions shot through with political context further illustrates their potential
for harm. It is hypothesized that the persistence of binaries as illustrated in this article could be politically motivated. Conventional spatial data producers such as national mapping agencies or cartographic companies may have economic or geopolitical reasons for encouraging the bifurcation of official and citizen-based approaches, in addition to the obvious motivation of retaining authority. Conversely, proponents of citizen spatial data production might also be politically motivated to encourage these binaries in order to advance, for example, anti-state ideologies or grassroots movements. Nonetheless, moving beyond this simplistic framework might better enable the cross-fertilization of approaches and the cooperation of diverse spatial data producers—be they employees of national mapping agencies, corporate mapping companies, or average citizens and Web users. Similarly, Johnson and Sieber (2013) noted that terminology can strongly influence whether citizen-generated data are used for official purposes; thus, labelling volunteer efforts as necessarily ‘amateur’ or ‘asserted’ “may serve to marginalize VGI as a data type, regardless of its fitness relative to authoritative sources” (72). Reconceptualizing spatial data production not as a binary but as a continuum could foster hybridity and pluralist engagements, and perhaps, the emergence of entirely new approaches to spatial data production. Doing so has a precedent—namely, the trajectory of critical GIS. Despite its initial reliance on a number of binary oppositions originating in the GIS and Society era, the longer-term maturation and success of critical GIS might be attributed to its decreased reliance on binaries and opposition, and increasing engagement with plurality and diversity.

Hybrid approaches to spatial data production hold the promise of greater availability of data through harnessing the strengths of each approach: the local knowledge, affordability, and speed of citizen-generated data production, and the oversight, organization, and quality control of more conventional approaches. Although this article contends that moving beyond the identified binaries could foster hybridity and plurality with respect to data production, this opportunity also poses a number of challenges. Managing spatial data produced by numerous, diverse, and often distributed contributors will require further attention to several key issues. In the case of datasets populated by diverse contributors, a number of questions must be addressed. Who owns datasets produced by diverse contributors? What end-users and organizations can access the data? For what purposes can it be used? What rules of governance apply? These questions are certainly more difficult to answer in the case of hybrid-produced datasets in comparison to conventional datasets. Scassa (2013) discusses a number of legal issues pertaining to VGI that would also pertain to hybrid datasets, including issues of intellectual property, liability, defamation, and privacy. With regard to intellectual property, it may be challenging to determine ownership of works or products derived from hybrid datasets produced by numerous contributors and entities. If intellectual property rules are not clear, individual contributors to a hybrid dataset could potentially claim ownership to products created from it. Similar challenges are posed by hybrid datasets with respect to liability and defamation concerns. If, for example, a hybrid spatial dataset contained incorrect information or was employed to defame an individual, it may not be apparent who is legally responsible. In terms of privacy, many datasets are subject to restrictions that determine how the data can be visualized, shared, and disseminated. These regulations are often clear-cut in the case of government-produced datasets for example, but in the case of hybrid government-citizen produced datasets,
privacy rules may be less obvious. Beyond legal issues, operational challenges are also evident, yet as described above, insights from STS and SDI research among other bodies of knowledge, will surely provide a foundation for a new wave of research on enabling hybridity between citizen and official data. Despite the significant challenges of hybridity, there exists great potential to unite citizen, government, and corporate spatial data producers. Perhaps the first step, however, is to move beyond the simplistic, binary conceptualization of contemporary spatial data production, and realize the potential that can be achieved by thinking outside of the box(es).

Acknowledgements

Thanks to Nadine Schuurman for the invitation to participate in the Alt.GIS meeting in 2013 and for helpful advice for approaching this research, and to the three anonymous reviewers for the insightful comments that have strengthened the paper. This work was made possible through funding from the Social Sciences and Humanities Research Council (SSHRC) for doctoral and post-doctoral studies.

References


\[1\] For an interesting discussion of the GeoWeb-GIScience relationship, see the transcript of the #geowebchat Twitter discussion from 6 March 2012, “Geoweb vs GIScience: two tribes or a field in tension?” at: http://mappingmashups.net/2012/03/06/geowebchat-transcript-6-march-2012/

\[2\] Interface can be found at: http://navigator.er.usgs.gov/.