Journal of Urban Technology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713436614

Visualizing Global Cyberscapes: Mapping User-Generated Placemarks

Mark Graham; Matthew Zook

Online publication date: 27 May 2011

To cite this Article Graham, Mark and Zook, Matthew(2011) 'Visualizing Global Cyberscapes: Mapping User-Generated Placemarks', Journal of Urban Technology, 18: 1, 115 — 132
To link to this Article: DOI: 10.1080/10630732.2011.578412
URL: http://dx.doi.org/10.1080/10630732.2011.578412

Journal of Urban Technology, Vol. 18, No. 1, January 2011, 115–132

Visualizing Global Cyberscapes: Mapping User-Generated Placemarks

Mark Graham and Matthew Zook

ABSTRACT *This article focuses on the representation of physical places on the Internet or what we term cyberscape. While there is a wide range of online place-related information available, this project uses the metric of the number of user-generated Google Maps placemarks containing specific keywords in locations worldwide. After setting out the methods behind this research, this article provides a cartographic analysis of these cyberscapes and examines how they inform us about the material world. Visibility and invisibility in material space are increasingly being defined by prominence, ranking, and presence on the Internet, and Google has positioned itself as a highly authoritative source of online spatial information. As such, any distinct spatial patterns within uploaded information have the potential to become real and reinforced as Google is relied upon as a mirror of the offline world.*

Introduction

The Internet surrounds us like air, saturating our offices and our homes. But it's not confined to the ether. You can touch it. You can map it. And you can photograph it (Blum, 2009).

In early November 2008 Google, the world's most used search engine, introduced a new means of highlighting the nexus between content on the Internet and the material world. This took the form of Google Flu Trends: a service that aims to improve people's health and even save lives by aggregating information collected from the searches of millions of Internet users. The website analyzes the location (at the scale of U.S. states) and timing of user-generated queries related to flu symptoms.[1] In doing so, Google claims Flu Trends functions as an early warning system that can detect where outbreaks are spreading (Ginsberg et al., 2008). It turns out that not only are Google's predictions strongly correlated with data collected by the Centers for Disease Control and Prevention, but most significantly are available seven to ten days earlier. This is because users are able to conduct informational searches immediately upon become symptomatic versus a more delayed visit to a clinic or doctor's office (Helft, 2008).

Although it is easy to discount Flu Trends as just another interesting, but largely insignificant use of the Internet, this paper argues that its public launch highlights an important shift in how people perceive the relationships between content available on the Internet and material space. By launching Flu Trends, Google is attempting to harness a collective intelligence present on the Internet to infer characteristics about physical corporeal space (indeed, few things are

1063-0732 Print/1466-1853 Online.

DOI: 10.1080/10630732.2011.578412

more material and non-virtual than an illness which has the potential to kill tens of thousands of people).

The existence of Flu Trends and the very fact that its stated mission is to use the keyboard strokes and mouse clicks of people sitting at computers to save actual lives leads to a number of crucial questions. What do the aggregate activities of people using and creating content on the Internet tell us about the material, social, and physical world? How strong are the linkages between representations of the physical world on the Internet and the very places those constructions represent? How can these connections be identified and studied by urbanists and social scientists? While Google Flu Trends illuminates an important connection between the material and the virtual, it is an extremely narrow example of the much larger array of human activity. This paper leverages another Google service (Google Maps) to outline some of the broad contours of the emerging ways in which virtual information intersects with everyday life.

Cyberscapes and Digiplaces

A map says to you, "Read me carefully, follow me closely, doubt me not. . .I am the earth in the palm of your hand." (Harley, 1989: 1)

Places are shifting, conflicting, and intersecting texts, and the ways in which they are represented have always been the subject of power struggles (e.g., Karimbayeva, 2010; Mels, 2006; Raento and Watson, 2000; Springer, 1985; Zook and Graham, 2007a). Any spatial representation (e.g., online spatial databases such as Google Maps) "stabilizes a particular meaning within a world of possible meanings. And in this modern world it generally does this by asking us to look at this thing, this object, this place" (Pickles, 2004: 3). By being abstractions from concrete realities, representations of the material world potentially facilitate the domination and control over the subjects of any representation (Crampton and Krygier, 2006). Barrow (2003), for instance, has detailed the ways in which colonial era maps of India were used to make British rule appear natural. Stickler (1990) has similarly documented the manner in which black settlements were often made invisible in maps of South Africa during apartheid.

The ability to map and represent material places has undergone a radical transformation in recent years. The practices of Neogeography (individuals creating and sharing maps via online tools, see Turner, 2006) and cloud collaboration potentially allow anyone with Internet access to contribute to the virtual layers of the palimpsests[2] of place (Graham, 2010b). These representations become part of our cognition of the places that surround us, that we move through, that we touch, see, and hear. It has long been known that people orient themselves based on mental maps (Lynch, 1960), and digital and online palimpsests now undoubtedly have become an important shaper of many people's mental maps.

Place can be represented in myriad ways online, but it is Web 2.0 mapping services that have really resulted in an explosion of place-based peer-production. This phenomena has been pointed to by a variety of authors and the ability for users to create and share spatial annotations has been variably termed volunteered geographic information (Goodchild, 2007), Maps 2.0 (Crampton, 2009), and Neogeography (Turner, 2006). But regardless of the specific labels used, the scale of this peer-production is enormous (Graham, 2010b). Just to name a few

examples, there are now about half a million geotagged places that have been contributed to Wikipedia, ten million that have been contributed to WikiMapia, eight hundred million GPS points uploaded to OpenStreetMap and almost ten million placemarks uploaded to Google (Graham, 2010a; Scott, 2009).

These millions of representations, all tagged to a particular material place on the Earth's surface, enter into the palimpsests of place that shape how we interpret, and therefore ultimately interact with the world (e.g., Parks, 2009). The fact that millions of people are creating many more millions of spatial representations (in addition to the existing listings in Google's directory) means that this emerging new layer of place is difficult to describe, map, and study. Yet, the importance of this new digital layer to our understandings of place means that new methods have to be adopted in order to answer some fundamental questions. Notably, what kind of information is being created about place? how does this information reflect the offline world it purports to represent? and most crucially, what places are being annotated? This paper sets out a methodology for approaching this final question and provides a preliminary look at the answer.

Before focusing on the methodology used to conduct this study, it is necessary to introduce two terms, *cyberscape* and *digiplace*, that offer a useful way to conceptualize the digital dimensions to material places that are being created. Cyberscape refers to the type, amount, and quality of geo-coded data on the Internet about places (Crutcher and Zook, 2009) . This virtual shadow to material places can consist of Google placemarks, Wikipedia articles, geotagged Flickr images, and any other information about a specific place that can be accessed non-proximately from outside of that place. Cyberscapes, in other words, are the virtual palimpsests that come into being through the mirroring of material places in networked documents, images, videos, and immersive environments.

Digiplace on the other hand refers to the sense of place created when cyberscapes influence our understandings and interactions with and through material place. Following Julie Cohen's formulation of the cyberspace metaphor, we use digiplace to refer to the extensions of everyday spatial practices: "an experienced spatiality mediated by embodied human cognition [. . .that] is relative, mutable, and constituted via the interactions among practice, conceptualization, and representation" (Cohen, 2007). Digiplaces are further fundamentally shaped by the ways in which dimensions of cyberscapes are made visible or invisible through networked electronic devices and their operating algorithms (Zook and Graham, 2007a; Zook and Graham, 2007b; Zook and Graham, 2007c). For example, methods of determining online visibility and invisibility such as search algorithms that rank millions of pages (e.g., a Google Maps ranking of a search on restaurants via a mobile device) and debates about how best to represent place in a wiki all influence the elements of the material world that get rolled into experienced digiplaces.

Together, cyberscapes and digiplaces shape the increasing hybridized spaces of human (particularly urban) activity. Neither completely virtual nor completely material, these cyber-spatial phenomena represent a new element in how people's views of the world are formed. Due to their newness, however, the contours and effects of cyberscape and digiplace remain relatively unknown and unstudied. This paper represents an initial effort to understand these events, primarily through the mapping of a range of cyberscapes, and set out a research agenda for exploring the issues raised.

Building and Exploring Cyberscape Data

This paper specifically analyzes the geographies of user-created content that is available via Google Maps placemarks, i.e., user created content that annotates a specific point on the earth. While this represents only one avenue through which people create and share spatially referenced content on the Internet, Google Maps has been selected as the object of study for two main reasons. First, Google plays a dominant role in the construction of representations about place. Studies (e.g., comScore, 2009) have shown that over fifty percent of people in most countries use Google to search for information (in many countries over ninety percent of people use the search engine).

Second, the company is also a leader in online mapping services, operating a variety of services, most notably Google Earth and Google Maps. Indeed, Google Earth has been downloaded 500 million times and almost ten million placemarks[3] have been created by users and are indexed by the company (Scott, 2009). Google has also been instrumental in the promotion of the Keyhole Markup Language (KML) which is the standard of the Open Geospatial Consortium, and has now become a de facto industry and consumer standard (Shankland, 2008). Created by Keyhole, Inc. (acquired by Google in 2004) KML files store textual descriptions, images, polygons, models, and of course geographic coordinates. Part of its popularity lies in the fact that it is easy to use (via stand alone software like GoogleEarth or online applications such as Google MyMaps) and view via a range of popular web applications and stand-alone software packages. As a result, Google has become a key (if not the key) node on the Internet for the creation and sharing of user generated content that references (via decimal latitude/longitude coordinates) specifics points on the earth.

Thus, Google Maps represents one of the best indexes of the Internet's "collective intelligence" about the material world. In other words it represents the global cyberscape, i.e., an aggregation of the comments, biases, passions, etc. that Internet users have towards specific places. To be sure, the demographics of Internet users have well known biases and the content creating a subset of the population may be even more skewed. Moreover, the notion of the "collective intelligence" of something as varied and decentralized as the Internet is somewhat problematic. Nevertheless, this paper argues that regardless of bias or agreement, the results generated by a keyword search at a specific location within Google Maps (or the digiplace created by searching the Google Maps cyberscape), are significant precisely because it is through this process that new digital palimpsests of place are created. For better or worse the cyberscape indexed by Google Maps is an increasing key way in which many (if not most) Internet users form their perceptions about places.

Building the Database

In order to analyze the Google Maps cyberscape, this project utilizes an automated and customized script (Burke, 2002; Gibson and Erle, 2006) that conducts Google Map queries and collects the number of hits (or placemarks) resulting from each search.[4] This result provides an indication of the level of spatial data available about that location relative to other locations, i.e., the intensity of a place's geospatial data availability. Queries were conducted during 2008 and 2009 with the specific dates referenced in the subsequent analysis.

A Google Maps query contains a number of different variables. For the purpose of this project the four key variables are as follows:

- NEAR – the location on the earth surface on which the search is centered (e.g., 38 N, 84.5 W);
- Q or QUERY – a keyword term that Google will match to online spatial data (e.g., bowling alley or church);
- RADIUS – the radial distance around the NEAR location that the search will include (e.g., 12 miles); and
- MRT –the type of search to be conducted (e.g., the Google Maps Directory, comparable to online yellow pages, or only user created data indexed by Google Maps).

Figure 1 outlines the result of a search where NEAR = 38.0 N 84.5 W, RADIUS = 10 miles, QUERY = "1", and MRT = user generated content. The number of hits resulting from this search was 3,128 and is the key piece of data collected by the automated script.[5]

In order to provide a meticulous look at the distribution of user generated content worldwide, this project systematically varied the values for the four factors. The NEAR variable consisted of pairs of latitude-longitude coordinates from a 1/4 degree grid of all the land mass in the world (excluding Antarctica). It is roughly 260,000 points in total. The RADIUS variable is a sliding value based on the great circle distance to neighboring points in the grid pattern. It was important to adjust this value in order compensate for decreasing distance between longitudes as one moves from the equator to the poles. The radius measure ranged from 12.2 miles at the equator to 9.6 miles at 60 degrees North or South. The goal was to completely cover the earth's surface while minimizing overlap with adjacent

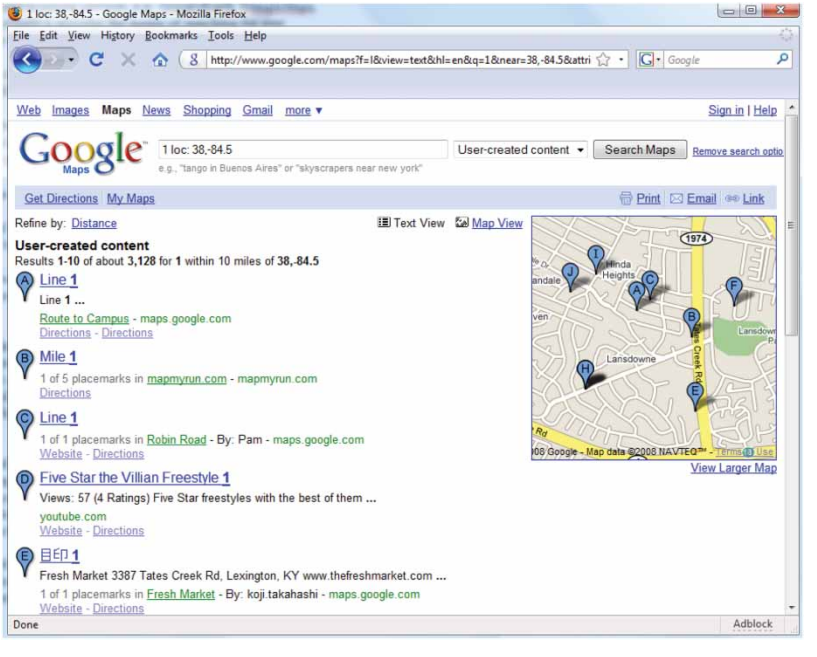

Figure 1. Example of Google Maps search results
Source: Author Screenshot

points—although this did occur since Google Maps search areas are circular and cannot be adjust to congruous geometric shapes (Christaller, 1933).

The QUERY variable is defined by a wide range of keywords (e.g., allah, sex, bookstore, etc.) and is essentially limitless although one must keep in mind potential linguistic bias. For example, searching for coffee would result in fewer results from non-English speaking locations (Germany or France); places that used non-ASCII character sets (Japan of China) or even places that have a cultural preference for tea (U.K.). Therefore keywords are selected that are as international as possible to reduce this bias.

Another important point is the need for a defined QUERY term to conduct a Google Maps query. Ideally one could search via a wildcard character to obtain a listing of all spatially referenced content but currently Google Maps does not allow this. Instead, this paper uses the keyword "1" (the number one) as a proxy for the overall amount of user generated content at a particular point. The assumption underlying this approach is that the distribution of placemarks containing "1" is effectively random and will not be unduly biased by linguistic or culturally differences. The final variable MRT is either restricted to user generated content or official Google directory listings. Both versions provide important insight on the dimensions of emerging cyberscapes although the latter builds upon existing data such as yellow page listings.

The results of each unique search, i.e., for a keyword at a specific point (defined by a set of latitude-longitude coordinates) using a specific search radius and limited to user generated content, is collected. For example, the entry for the unique search represented in Figure 1 would be 3,128 as this was the number of placemarks identified.

Goals for Exploring the Database

The goal of this paper is to explore the global cyberscape along two dimensions. First, building upon previous work that documented the uneven geography of Internet content creation (Gorman and Malecki, 2002; Zook, 2005a; Zook, 2005b), it is useful to examine the extent to which this variegation persists within volunteered geographic information. Therefore the density of user-created content worldwide is examined at a range of scales. Given the historical precedents, is it is expected that user-created representations of place will display similarly uneven patterns.

The second goal of this paper is to explore the distinct characteristics of cyberscapes via mapping the contours of the use of specific keywords. Visualizations of the characteristics of cyberscapes will ultimately allow the links between the material and virtual domains to be more fully examined: links that ultimately become part of the hybrid places we experience and interact with. The final objective of this paper is outlining a research agenda for exploring these two dimensions of cyberscapes in much more detail with the additional step of understanding how they affect our cognition of places.

The Uneven Geography of Global Cyberscapes

As users annotate the millions of placemarks that now blanket the Earth, the information that can be accessed about places is changing. Visibility and invisibility in physical space are increasingly being defined by prominence, ranking, and

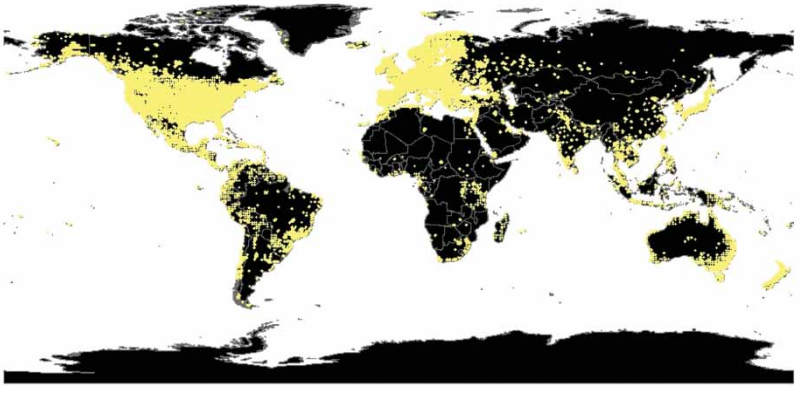

Figure 2. Points with more than 100 hits, July 2008
Source: Author Survey

presence in online information. It is conceivable that distinct patterns within cyberscapes have the potential to become real and reinforced as online information is relied upon to mirror the offline world. It is, therefore, useful to first examine the geography of user-created content as a whole. What places are being annotated? How does this vary by scale and topic? And do representations on the Internet offer an accurate mirror of the underlying material world?

Figure 2 demonstrates that peer-produced information is far from evenly distributed at a global level.[6] Certain parts of the world are covered by a dense cloud of virtual information, while most of the planet has only a small amount of online representation. Interestingly, the geography of peer-produced information is not uneven due to the global unevenness in population density. (See Figure 3.) Indeed, there are enormous disparities when these data are examined on a finer scale. For instance, Louisville Kentucky has almost fifty percent more user-created content about it than the entire country of Iraq. Even more astonishingly, the Tokyo metropolitan region is represented by three times as much content as the entire continent of Africa.

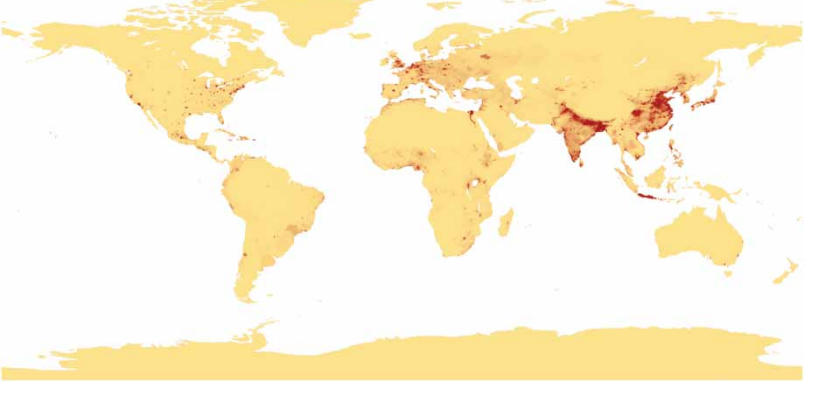

Figure 3. Population density
Source: http://visibleearth.nasa.gov/view_rec.php?id=116

Figure 4. Placemarks in China, July 2008
Source: Author Survey

Information inequalities can be observed at all scales in every part of the globe, but Figure 4 offers a particularly striking example. This map displays the density of placemarks in China. Beijing, Shanghai, and the Pearl River Delta region all are characterized by heavy information densities, while the rest of China is represented by only a small amount of information[7]. These data indicate that, at a meta-level, cyberscapes are at their most vivid and rich in only certain clusters of high-density virtual information.

Although this unevenness is not unexpected, it is important to highlight this fact prior to undertaking a more nuanced mapping utilizing a range of keywords. Can cyberscapes illuminate more subtle differences between places or is it limited to a crude distinction between more connected and less connected places?

Keyword Contours of Cyberscape

As illustrated by the example of Google's Flu Trends in the introduction of this paper, cyberscapes can potentially mirror many aspects of the material world. To more closely explore the ways in which cyberscapes link back to observable characteristics of the material world, a number of keywords (including both user generated and Google directory listings) are presented in the following analysis of more specialized dimensions of cyberscapes.

Business and Sex in the Global Context

With this breadth in mind, the keywords "business" and "sex" are used as globally used terms that represent significant memes within the globalization process. In Figures 5 and 6, the size of the black circles indicates the absolute number of references to either "business" or "sex" in user-created Google placemarks. The shading of each map represents the specialization in references to each term (each term was compared to an index of all other user-generated content).

Sex and business clearly have distinct albeit related geographies. Not surprisingly the Global North has the largest concentration of placemarks: a fact consistent with the information inequality noted earlier in the paper.

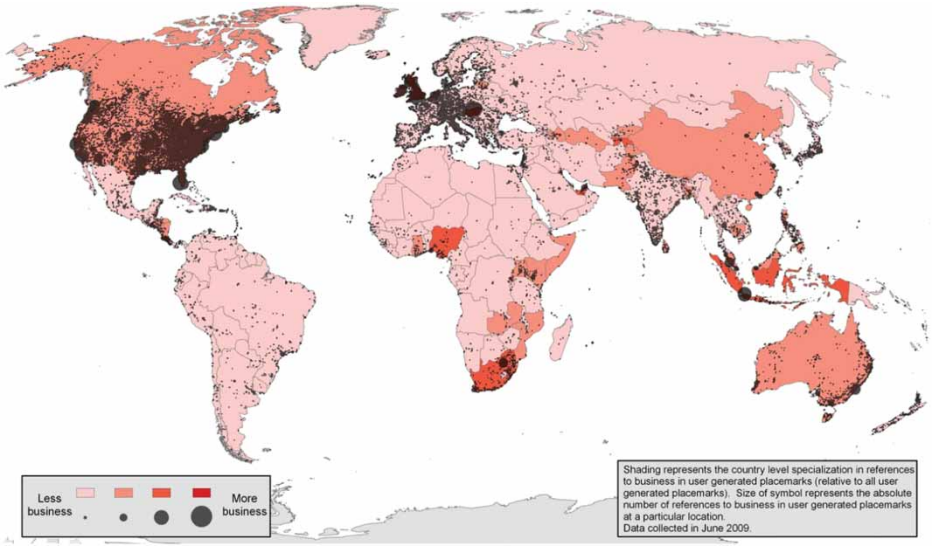

Figure 5. User-generated references to business in Placemarks, June 2009
Source: Author Survey

North America, Japan and much of Europe are largely blanketed by references to business, while most of the rest of the world is characterized by far fewer virtual references. The UK and North America also have a high degree of specialization in terms of references to business, but high values are also present in non-Western countries that have strong ties to global business networks. As the largest low cost manufacturer, China shows a high degree of business specialization as does much of Central America which recently entered

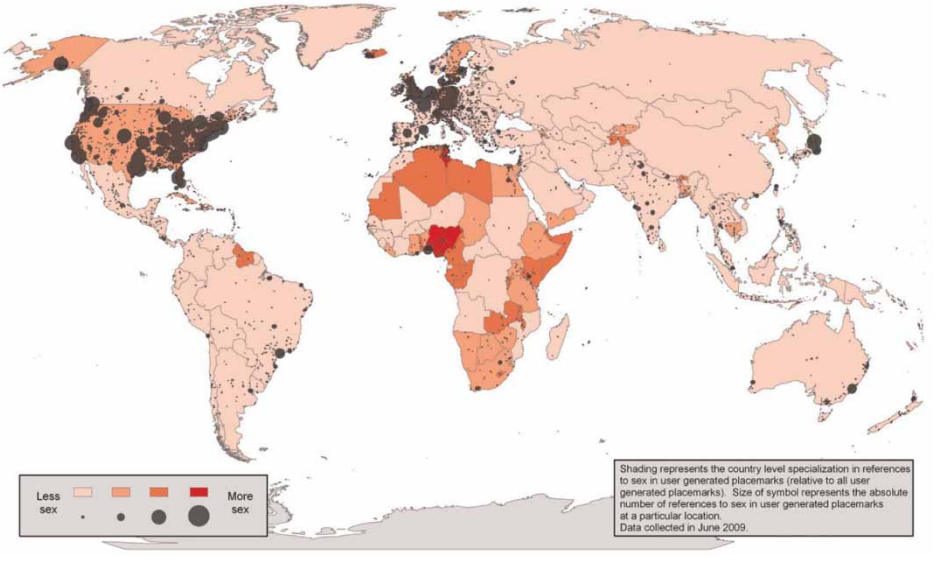

Figure 6. User-Generated references to sex in Placemarks, June 2009
Source: Author Survey

the Central America Free Trade Agreement (CAFTA) accord. The two largest economies of sub-Saharan Africa (Nigeria and South Africa) are specialized in business, as is the U.A.E. Other countries such as Indonesia and Hungary are highly specialized as well.

Interestingly, references to business are much more geographically dispersed than references to sex. Again, in absolute terms, the United States, Northern Europe and Japan have by far the most references to sex. However, when looking at specialization, intriguing patterns emerge. The United States and parts of Northern Europe (particularly the U.K., Sweden, Germany, the Netherlands, and Iceland) continue to be ranked highly.

Yet it is large parts of Africa that contain the highest degree of specialization. Or, in other words, user-generated content in countries like Nigeria, Kenya and Tunisia is far more likely to contain references to sex than user-generated content in most other places. While one would expect to see a degree of specialization in countries like the Netherlands (due to the well known sex industry in Amsterdam), the amount of specialization in places like Mauritania, Zambia and Lesotho is surprising. It could simply be a spurious result based on the generally low number of user-generated placemarks in those locations. Alternatively it suggests that "sex" may be one of the first topics in which people comment about a place, and it is only later that other foci appear.

Bibles and Bibliophiles

Moving from the global to the national scale, it is possible to compare the cyberscapes of religion and reason within North America. Based upon the number of Google Maps directory listings[8] for "churches" and "bookstores," Figure 7 illustrates an intriguing depiction of the geographic distribution of attitudes on reli-

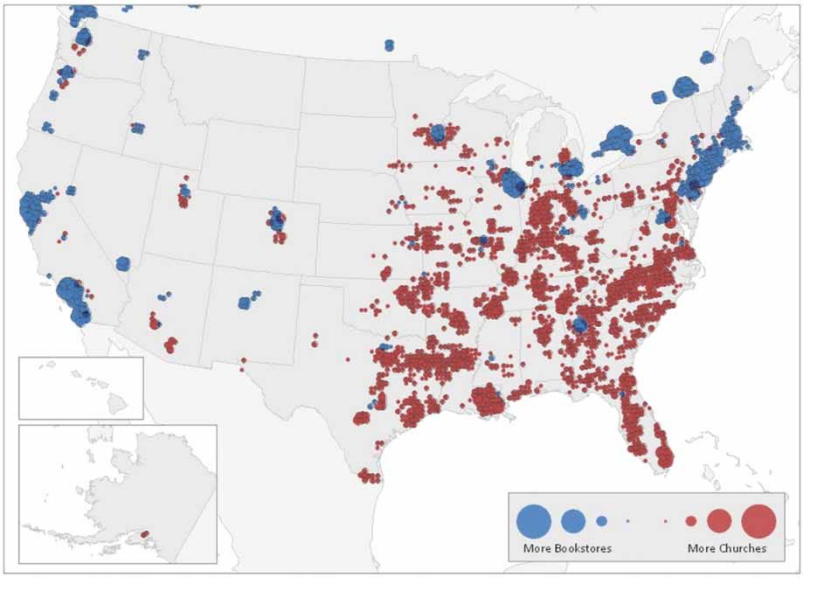

Figure 7. Prevalence of bookstores versus churches in Google Directory listings, August 2008
Source: Author Survey

gion and secularism within American society. This depiction is based not on surveys of individuals but drawn from online data about the resources available at each location, i.e., how many churches and how many bookstores.

As there are an overwhelmingly larger number of churches than bookstores nationwide it is necessary to index each variable before comparison. The technique used in Figure 7 was to divide the number of churches (or bookstores) at a location by the national average of churches or bookstores. If a location had twice the number of churches as the national average it would receive an indexed value of 2. Similarly having only 50 percent of the national average of bookstores would produce an indexed value of 0.5. The church index was then divided by the bookstore index to see each location's relative balance of churches to bookstores. If each of the indexed values were the same, the church-bookstore index would be equal to 1. But as in the case of the example above (church index = 2, bookstore index = 0.5) the final index would be 4. This indicates that this particular location has a much higher relative number of churches to bookstores. In order to highlight places that had a higher specialization in either churches or bookstores, Figure 7 only includes locations where the church-bookstore index was skewed more than 20 percent in either direction (i.e., values greater than 1.2).

For the most part, the relative prevalence of bookstores occurs in and around the big cities - Los Angeles, California is the site of the highest indexed value, and is joined by the megalopolis of the eastern seaboard as having the highest concentrations in favor of bookstores. Even cities such as Atlanta, nestled in the Bible Belt of the American southeast, tend towards a relatively large number of bookstores. On the converse, other large cities like Dallas, San Antonio and Houston continue to favor churches, with New Orleans having the highest relative concentration of churches in the nation. Suburban areas surrounding large population centers also show near-universal favoritism for churches.

So while there appears to be no single variable determining the local trends toward faith or reason, it is evident that even some of the most common assumptions regarding the geographies of faith and reason have proven to be more complicated; not all large cities are necessarily bookish, but neither is the Bible Belt a homogeneous geographic unit.

Church, Bowling, Guns, and Strip Clubs

Mapping out the Bible Belt in American cyberscapes can be likewise accomplished by comparing a variety of other search terms. Again using the number of listings indexed by the Google Maps directory, Figure 8 visualizes the comparative prevalence of churches, bowling alleys, guns, and strip clubs across North America.[9] Each point is color coded according to which activity had the highest number of hits in the Google Maps directory.

Upon first glance, it is easy to see the relative supremacy of two topics, churches (in blue) and guns (in green), which cover most of the points in North America. Churches dominate throughout most of the southeast and upper Midwest while the Northeast, the West and much of Canada show a higher number of listings for guns. One should not, however, interpret this to mean that guns are more prevalent in Canada than the southern United States which clearly does not equate with the offline reality. Instead, the Northeast and Canada have relatively fewer listings for churches than in the South, leading

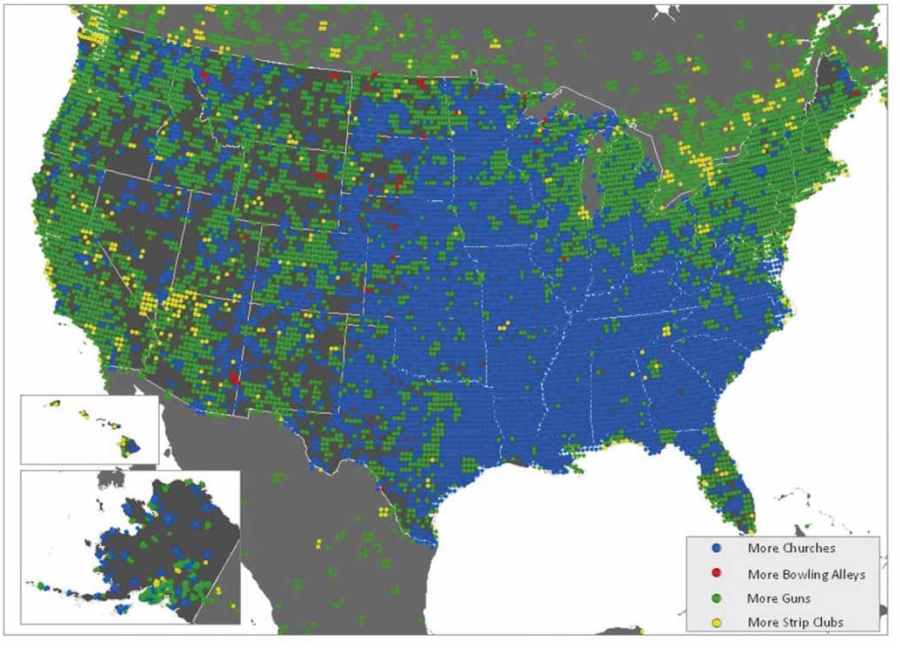

Figure 8. Prevalence of churches, bowling alleys, guns and strip clubs in Google Directory listings, August 2008
Source: Author Survey

them to be color coded with the next leading activity, i.e., guns. Likewise, there are plenty of gun listings in the southeastern United States; they are simply overshadowed by the listings for churches. Bowling alleys and strip clubs, in contrast, are far less visible in American cyberscapes outside of relatively small, specialized pockets (such as Las Vegas and Toronto).

The central point of these maps of North America is that an analysis of cyberscapes provides insights into cultural, political and economic trends in ways that have never before been methodologically possible. In Figures 7 and 8, The Bible Belt is no longer just straightforwardly associated with the entire American southeast. But rather, more nuanced visualizations are possible depending on the specific keywords and comparison words chosen.

User Generated Geographies of Religion

Returning to the global level of analysis, this paper provides a similarly complex illustration of the distribution of religion that traditional representations like Figure 9 gloss over. In this map, large swathes of the world are shaded to represent dominant religious groups within those areas, and the lack of a consistently fine-scaled, global-level survey of religious beliefs makes such an approach necessary. But, this homogenous representation ignores the very real on-the-ground heterogeneity of religious practice that exists.

While it is tempting to view maps of cyberscapes as simply updated versions of traditional maps such as Figure 9, the level of detail and range of topics available with cyberscape data represents a real innovation. In short, cyberscapes can convey aspects of the material world undetectable by other methods by accessing

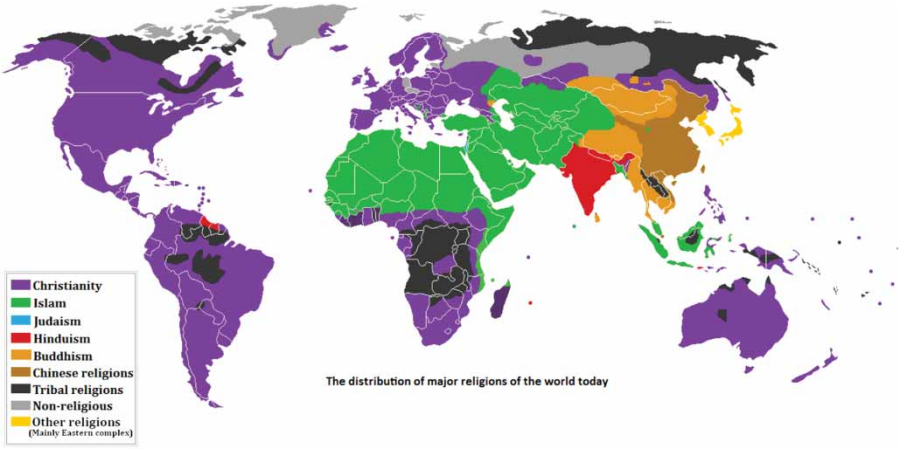

Figure 9. Religions of the world
Source: en.wikipedia.org/wiki/File:Religion_distribution.png

the global layer of digital representations based on cyberscapes and our lived and experienced digiplaces. So, instead of the United States being shaded as a consistent tone of blue to represent Christianity or Morocco being shaded as a consistent tone of green to represent Islam, Figure 10 illustrates a finer grain mapping of religion. Figure 10 visualizes the comparative prevalence of the terms "Allah," "Buddha," "Hindu," and "Jesus" (chosen due to their linkages with the major religions of Islam, Hinduism, and Christianity). Each point is color coded according to which term had the most user-generated references within Google Maps at a particular location.

In this map there are distinct geographies of religious terms in user-created content indexed by Google. It can be seen that high rankings (in the number of

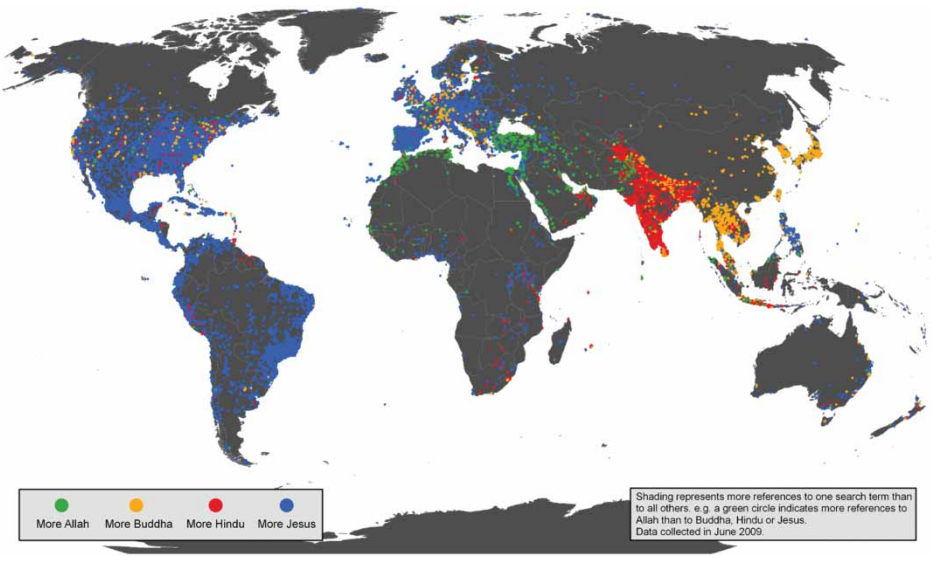

Figure 10. User-generated references to religion in Placemarks (GLOBAL), June 2009
Source: Author Survey

references) are often found in the most likely regions. For example, the Middle East, North Africa, and Muslim parts of South and Southeast Asia are all characterized by the largest number of references to "Allah." The largest numbers of references to "Buddha" are similarly clustered in East and Southeast Asia, the Himalayas, and Sri Lanka. The geography of references to "Hindu" is even more clustered. Here, the Indian Subcontinent, Afghanistan, Angkor Wat, Bali, Singapore and Kuala Lumpur (two cities with large Indian populations) have a large number of references. References to "Jesus" are more broadly distributed than any of the other three terms, but still show an incredible degree of concentration. The Americas, Western Europe and the Philippines are blanketed by references to Jesus.

Yet in very few parts of the world is there any semblance of a homogenous block of religious references. For example, as Figure 11 illustrates, the Indian subcontinent, which was represented as Hindu in Figure 9 there remain more references to Hindu than any of the other three search terms. But it can also be seen that there are parts of the country in which virtual references to "Allah," (in the Kashmir region) or "Buddha," (in the northern Himalayas as well as Sri Lanka) are more prevalent than references to "Hindu." Likewise one sees spots within the United Arab Emirates that are tagged "Hindu" despite the official dominance of Islam which likely reflects the presence of many expatriates from the Indian subcontinent who are working there. Thus, despite its many biases and significant blank spots (for example the much richer layer of information covering the Global North than the Global South), Figures 10 and 11 represent not just some of the complicated offline geographies of religion, but also the virtual cyberscapes that become parts of the experienced sense of place for millions of people around the globe.

It should again be stressed that this method simply provides us with a quantitative count of the number of references to each term in any given location. It is,

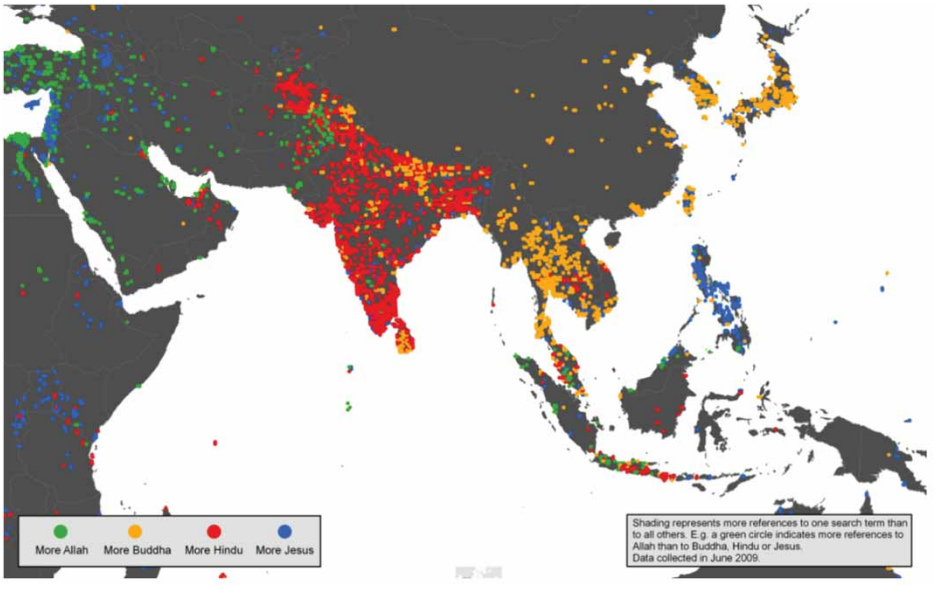

Figure 11. User-generated references to religion in Placemarks (ASIA), June 2009
Source: Author Survey

therefore, important to interpret these results with caution. A large number of hits for any term in a certain place may not necessarily reflect a large number of offline/material instances of that term. For example, a region dominated by one mega-church may have more references to "church" or "Jesus" than another region containing many churches. The method does however still give a useful indication of the amount of content referencing certain ideas, terms and practices in any part of the world. Nonetheless, groundtruthing and in-depth qualitative interpretations remain necessary if further claims are to be made about the nature of geographic representations in any specific place.

Setting Out a Research Agenda

This paper has demonstrated two main facets of the peer-production of cyberscapes. First, cyberscapes do not form a cloud that evenly covers all parts of the planet. The cloud of virtual information superimposed over place is thick and dense over some parts of the world, and little more than a wisp over others. In particular, much of Africa and large parts of Asia and Central America are represented by only thinly layered cyberscapes, if any at all.

These presences and absences play a fundamental role in shaping the ways that we interpret and interact with the world. The fact that the geographies of cyberspace content are so uneven therefore leads to worrying conclusions. As we increasingly rely on peer produced information, large parts of the world remain "terra incognita" (in a similar manner to the ways in which many of those same places were represented on European maps before the 19th Century). Moreover, the little representation that exists in peer produced cyberscapes, is often created from outside the developing world; a phenomenon that can be observed in content produced by other practices of neogeography (Cohen, 2006; Graham, 2009).

These uneven layers of representation are undoubtedly influenced by both the uneven geographies of Internet infrastructure and a range of other social, economic, and political factors (Dodge and Kitchin, 2001; Gorman and Malecki, 2002; Zook, 2001). Censorship and suppression in particular have the potential to shape not only the amount of representation about certain places, but also the content and subject of those representations (Zook and Graham, 2007a). The costs and availability of bandwidth are also an especially important factor: serving to exclude the contributions of people from many parts of the world with limited broadband internet access (Dutta and Mia, 2010). Conversely, parts of the world characterized by a large amount of virtual content are characterized by information clutter, thus giving power to the sorting systems and algorithms that determine which aspects of place are more visible than others, *i.e.*, the shape of digiplaces. Future research should ultimately focus on addressing these many links that exist between offline/material factors and the production of Internet content.

The second finding is that despite their unevenness, cyberscapes allow an unprecedented insight into aspects of the material world that would otherwise be practically impossible to measure. The analysis of religious cyberscapes, for example, reveals patterns at a scale and scope that is practically impossible to observe with more conventional methods. However, it should be recognized that due to the geographic biases inherent in both the subject and the production of geotagged information, cyberscapes can only ever serve as distorted mirrors of the features, trends, and characteristics that they reflect.

This paper has described the first step and background of a research agenda that ultimately seeks to understand the two facets of cyberscapes in much more detail. The methodology outlined in this article provides an initial analysis. There remain, however, many questions (both technical and theoretical) to be asked including:

- Which places, and which aspects of those places, are being mapped?
- Who is writing this information?
- How accurate/reliable is this information?
- How is it filtered, ordered and ranked?
- In what ways can it be measured?
- How will cyberscapes and digiplace change the way we see, move through, and use space?

As search engines like Google start to become the de facto sources of global knowledge, it is crucial to understand and map out the politics and biases embedded into the algorithms that order information, the geographies of peer-produced information, and the content of that information (S. Graham, 2005; Hargittai, 2007; Introna and Nissenbaum, 2000). The paper has taken a first step towards understanding some of these issues and has outlined a methodology for mapping the distorted mirrors of cyberscapes that ultimately have the potential to become real and reinforced as they become integrated into our lived experiences of place.

Notes

1. Although this paper focuses on Google, it is not the only organization to use the collective intelligence of cyberspace to make inferences about economic, social, and political relationships in physical space; it is simply the most visible. Yahoo is working on a similar service, as are smaller groups such as Whoissick.org and HealthMap.
2. The idea of a palimpsest refers to the myriad material, historical, and virtual layers that shape our understanding of a place. Palimpsests, therefore, consist of not only material experiences of place, but also photographs, videos, stories, websites, and countless other objects and representations.
3. This statistic was calculated by using Google to search for the following terms “filetype:kmz” and “filetype.kml.”
4. Also see *http://datamining.typepad.com/data_mining/* for an excellent overview of other datamining techniques.
5. In comparison, the same search but for all of Google indexed data returned 10,273 hits.
6. As it was not possible to obtain the total number of placemarks at particular location when this data was collected, this paper uses the keyword "1" (the number one) as a proxy for the overall amount of user generated content at a particular point.
7. Interestingly, Karimbayeva (2010) has found that the ability for users to contribute geographic information to Google from within China is often limited. It is therefore possible that many of these contributions are submitted by users from outside of the country.
8. Google Maps directories are drawn from a range of sources such as yellow page listings and appear in Google Maps search results as individual placemarks. These placemarks are distinct from and exclude user generated placemarks.
9. These four terms are selected due to the fact that they represent a variety of ways in which people can spend their free time.

Bibliography

I.J. Barrow, *Making History, Drawing Territory: British Mapping in India, c. 1756–1905* (New Delhi: Oxford University Press, 2003).

A. Blum, "Netscapes: Tracing the Journey of a Single Bit," *Wired* (2009) <http://www.wired.com/magazine/ff_internetplaces> Accessed August 06, 2010.

S.M. Burke, *Perl & LWP* (Sepastopol, CA: O'Reilly Media, 2002).

W. Christaller, *Die Zentralen Orte in Suddeutschland* (Jena: Gustav Fischer, 1933), translated (in part) by C.W. Baskin, as *Central Places in Southern Germany* (New York: Prentice Hall, 1966).

J.E. Cohen, "Cyberspace as/and Space," *Columbia Law Review* 107 (2007) 210–256.

N. Cohen. "African Languages Grow as a Wikipedia Presence," *New York Times* (August 31, 2006) <http://www.nytimes.com/2006/08/31/business/media/31link.html> Accessed August 06, 2010.

comScore. "U.S. Search Engine Rankings," comScore (2009) <http://www.comscore.com/Press_Events/Press_Releases/2009/3/US_Search_Engine_Ranking> Accessed August 06, 2010.

J.W. Crampton, "Cartography: Maps 2.0," *Progress in Human Geography* 33:1 (2009) 91–100.

J.W. Crampton and J. Krygier, "An Introduction to Critical Cartography," *ACME: An International E-Journal for Critical Geographies* 4:1 (2006) 11–33.

M. Crutcher and M. Zook, "Placemarks and Waterlines: Racialized Cyberscapes in Post-Katrina Google Earth," *Geoforum* 40:4 (2009) 523–534.

M. Dodge and R. Kitch, *Mapping Cyberspace* (London: Routledge, 2001).

S. Dutta and I. Mia, *The Global Information Technology Report 2009–2010* (Davos: World Economic Forum, 2010).

R. Gibson and S. Erle, *Google Maps Hacks* (Sebastopol: O'Reilly Media, 2006).

J. Ginsberg, M.H. Mohebbi, R.S. Patel, L. Brammer, M. Smolinski, and L. Brilliant, "Detecting Influenza Epidemics Using Search Engine Query Data," *Nature* (November 19, 2008) <http://www.nature.com/nature/journal/vaop/ncurrent/full/nature07634.html> Accessed August 06, 2010.

M.F. Goodchild, "Citizens as Sensors: The World of Volunteered Geography," *GeoJournal* 69:4 (2007) 211–221.

S.P. Gorman and E.J. Malecki, "Fixed and Fluid: Stability and Change in the Geography of the Internet," *Telecommunications Review* 26:7–8 (2002) 389–413.

M. Graham, "Wikipedia's Known Unknowns," *Guardian* (December 2, 2009) <http://www.guardian.co.uk/technology/2009/dec/02/wikipedia-known-unknowns-geotagging-knowledge> Accessed August 06, 2010

M. Graham, ""Cloud Collaboration: Peer-Production and the Engineering of Cyberspace,"," in S. Brunn, ed., *Engineering Earth* (New York: Springer, 2010a).

M. Graham, "Neogeography and the Palimpsests of Place: Web 2.0 and the Construction of a Virtual Earth," *Tijdschrift voor Economische en Sociale Geografie* 101:4 (2010b) 422–436.

S.D.N. Graham, "Software-Sorted Geographies," *Progress in Human Geography* 29:5 (2005) 562–580.

E. Hargittai, "The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction," *Journal of Computer Mediated Communications* 12:3 (2007).

J.B. Harley, "Deconstructing the Map," *Cartographica* 26:2 (1989) 1–20.

M. Helft, "Google Uses Searches to Track Flu's Spread," *New York Times* (November 12, 2008)<http://www.nytimes.com/2008/11/12/technology/internet/12flu.html> Accessed August 06, 2010.

L.D. Introna and H. Nissenbaum, "Shaping the Web: Why the Politics of Search Engines Matters," *The Information Society* 16 (2000) 169–185.

Z. Karimbayeva, *Censoring Maps in Google China? Visual Analysis through Foucault's Power/Knowledge*, (Department of Geosciences: Georgia State University 2010).

K. Lynch, *The Image of the City* (Cambridge, MA: The MIT Press, 1960).

T. Mels, "The Low Countries' Connection: Landscape and the Struggle over Representation Around 1600," *Journal of Historical Geography* 32:4 (2006) 712–730.

L. Parks, "Digging into Google Earth: An Analysis of Crisis in Darfur," *Geoforum* 40:4 (2009) 535–545.

J. Pickles, *A History of Spaces* (London: Routledge, 2004).

P. Raento and C.J. Watson, "Gernika, Guernica, Guernica? Contested Meanings of a Basque Place," *Political Geography* 19:6 (2000) 707–736.

L. Scott, "Personal Correspondence with Laura Scott: Communications and Public Affairs Officer at Google" (2009).

S. Shankland "Google Mapping Spec Now an Industry Standard" *Cnet* (2008) <http://news.cnet.com/8301-10784_3-9917421-7.html> Accessed August 06, 2010.

C. Springer, "Textual Geography: The Role of the Reader in *Invisible Cities*," *Modern Language Studies* 15:4 (1985) 289–299.

P.J. Stickler, "Invisible Towns: A Case Study in the Cartography of South Africa," *GeoJournal* 22:3 (1990) 329–333.

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

M. Zook, "Old Hierarchies or New Networks of Centrality? The Global Geography of the Internet Content Market," *American Behavioral Scientist* 44:10 (2001) 1679–1696.

M. Zook, "The Geographies of the Internet," *Annual Review of Information Science and Technology* 40 (2005a) 53–78.

M. Zook, *The Geography Of The Internet Industry: Venture Capital, Dot-Coms, and Local Knowledge* (Malden, MA: Blackwell, 2005b).

M. Zook and M. Graham, "The Creative Reconstruction of the Internet: Google and the Privatization of Cyberspace and DigiPlace," *Geoforum* 38 (2007a) 1322–1343.

M. Zook and M. Graham, "From Cyberspace to DigiPlace: Visibility in an Age of Information and Mobility," in H.J. Miller, ed., *Societies and Cities in the Age of Instant Access* (London: Springer, 2007b).

M. Zook and M. Graham, "Mapping DigiPlace: Geocoded Internet Data and the Representation of Place," *Environment and Planning B: Planning and Design* 34:3 (2007c) 466–482.