The use of geovisualization to public health, in the context of open source applications and digital earths: an effective representation?

Ana Brandusescu,[1] Renee Sieber[2] and Nadine Schuurman[3]
1 Geography, Simon Fraser University, anab@sfu.ca
2 Geography/School of Environment, McGill University, renee.sieber@mcgill.ca
[3]Geograph, Simon Fraser University, Nadine@sfu.ca

Abstract

The use of geographic visualization (geovisualization) has increased over recent years with the prevalence of digital earths (e.g., Google Earth). We investigate how well applications of digital earths employ geovisualization principles using the example of health data. We utilized trauma data from Cape Town, South Africa. We created a number of geovisualizations at both the coarse level and fine grained level of resolution. These utilized various application programming interfaces (APIs), Google Earth, and scripting languages. We based our applications on evaluation of geovisualization design principles and extant examples in health and epidemiology.

Background and Relevance

The use of geographic visualization (geovisualization) has increased over recent years with the emergence of digital earths, especially the Google Earth (GE) platform (Dykes & Wood, 2007; Schultz et al., 2008; Butler, 2006). As their popularity increases, so do new ways of representing data from various fields. Applications such as mashups are starting to emerge in the health field (Chang et al., 2009; Zhang et al., 2009; Auchincloss & Roux, 2008; Sui & Holt, 2008; Janies et al., 2007). However, the investigation of how well health applications of digital earths employ geovisualization principles is relatively new.

According to MacEachren and others, the goal of geovisualization is to reveal the unknown and non-obvious and communicate it in an effective manner through end user interaction (MacEachren & Taylor, 1994; Kraak & MacEachren, 2001). There is some question whether these new platforms are sufficiently advanced in their capacity to adhere to these and other principles (e.g., in Cartwright et al. 2001) and certainly whether the digital earths are superior to more traditional GIS (Elwood 2009). Conversely, digital earths as part of the “read-write” web of bottom-up non-expert contributions, is compatible with MacEachren and Taylor’s (1994) vision, as opposed to the expert-driven nature of GIS. These have allowed us to discover ways of interacting with spatial data in a way that has not been implemented before.

This paper investigates geovisualization in digital earths as it pertains to trauma data. These data are useful as they are binary (one is either injured or not) and provide an example of spatial data clustering – which in turn permit analysis of risk associated with specific geographic areas like neighbourhoods. In addition, the Cape Town injury data set incorporated data from informal townships that are not included in the official census. This is a common phenomenon in low resource settings and often results in undercounting of the most vulnerable populations. These data – though not a complete population – cover three years and thus allow for temporal aggregation resulting in a better overview of patterns. For the purposes of this study, injury data support both spatial pattern recognition and the basis for higher resolution investigation of on-the-ground conditions based on vulnerability of specific neighbourhoods.

We primarily focus our attention on the ease of use of these new platforms for the developer and researcher. Our research question aims to answer the reasons for choosing specific geovisualization techniques such as two dimensional versus three dimensional, static images versus animation, coarse and fine resolution visualization, and repurposing of application add-ons to digital earths. We also evaluate domain-specific issues in the application of digital earths to the health field, such as confidentiality of data.

Data and Methods

The data sets utilized in this research all derived from Cape Town, South Africa: aggregate trauma injuries, suburbs and population (raw population and population density maps). The data is monthly and covers 6,644 injuries, out of which 4,269 (64%) can be used. The data is based on one of the two health catchment areas, which cover 174 suburbs in the city. To make the data compatible with the applications, extensive refinement of the data was required. This is typical of health data (Sack et al. 2009; Le, 2005). The research took place over sixteen months, (July 2007 to October 2008) and builds on prior work (Cinnamon and Schurmann, 2010).

To evaluate applications to be used in our research, we created a matrix of existing health and digital earth applications to better understand the relationships between geovisualization and health. Matrix criteria were formed from the geovisualization literature, including ease of modifiability and collaborative work, cost and other resource requirements, ability to integrate multimedia, quality of application user interfaces, and interactivity. We focused on the developer side of geovisualization so, for instance, we looked at interactivity in terms of the ease of the user's movement about the data as well as the interoperability of applications. Based on these criteria we developed a number of applications, which included a variety of multimedia, dimensional visualizations, and multiple geographic resolutions.

Results

We evaluated numerous digital earth applications for trauma data against those geovisualization design criteria. Certain applications, such as Jamstec (developed for the field of geology) were free but difficult to modify. Google Earth (GE) Graph was extensively used in creating many of the geovisualizations due to its high level of interoperability with GE. It also rated highly in its interoperability among different programs (e.g., GE, GE Graph, Google Docs, and Google Charts). Figure 1a, shows a 3D bar graph model of aggregate suburb injuries, using GE Graph.

Figure 1b shows data that was visualized within the GE 'information balloon'. The graph is generated "on the fly" from data in Google Spreadsheets (i.e., the chart is automatically updated in GE when it is changed in Google Charts). This relied on Google Charts's embedded application programming interface (API). Use of an application of Google Charts also offered opportunities for collaboration because multiple people can log in to modify the data in "the cloud". One way we chose to represent time or 4D was to use two bars per suburb and two graphs in an information balloon (not shown).

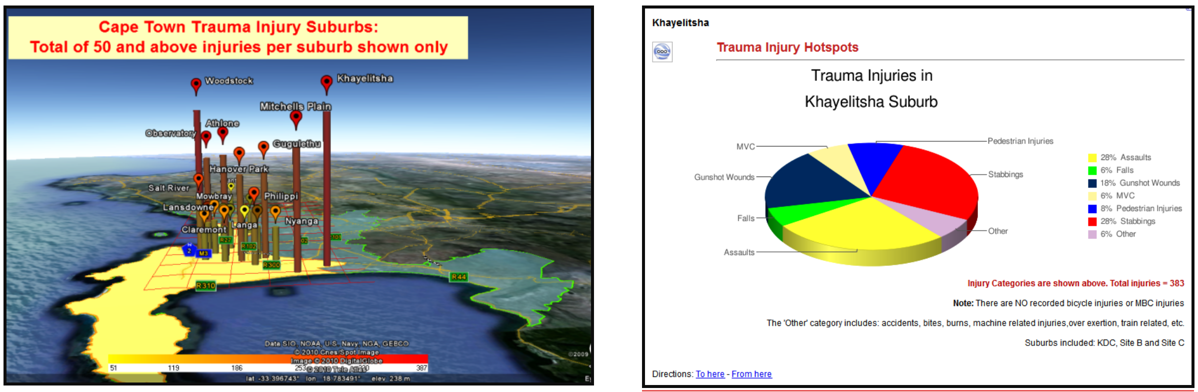

Figure 1.a. GE Graph model of Cape Town's top ten injury prone suburbs, and municipal catchment (green and yellow) areas, represented in Google Earth. 1.b Google Earth visibility balloon created with Google Charts API, representing trauma injury categories in Khayelitsha, Cape Town.

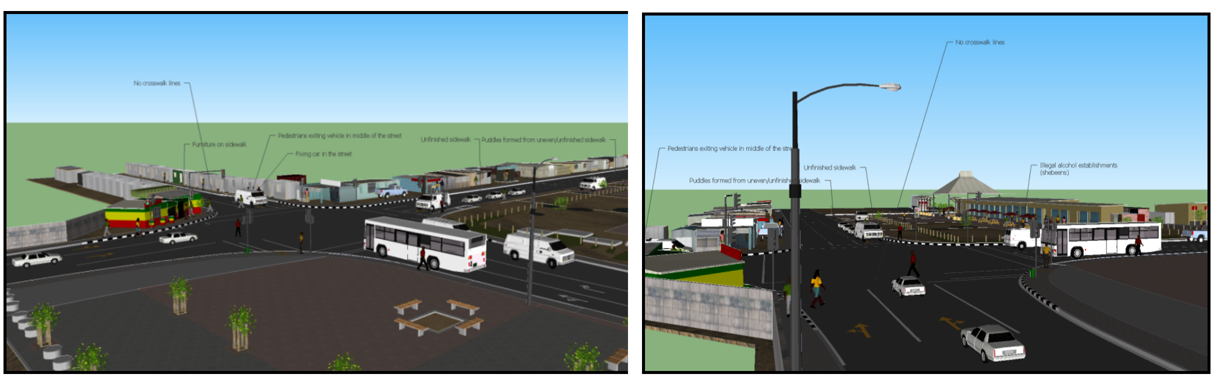

Figure 2.a Google Sketchup of a street intersection that causes numerous cases of trauma. 2.b Same Sketchup model but at a different angle.

Figure 2 shows our use of Google SketchUp to model a street intersection in Cape Town. The Sketch up model was used to visualize possible connections between environmental form and traumas. This includes places where the sidewalk is degraded allowing cars to come up on the pavement or forcing pedestrians into oncoming vehicles. The above views represent approximately 300 objects, which allowed for great visual depth but which frequently crashed laptops.

We have numerous findings. These applications, through incorporation of multimedia, dimensionality, and ability to pan, zoom and “walk” through the landscape, appear provide rich information for users who wish to explore and learn more about domain processes. The geovisualizations offered via the interfaces appear to assist in complex tasks. They can easily be customized; they need not be a one-size fits all approach (Dykes and Wood 2007). We argue that the difficulty in using traditional interfaces like those in ArcGIS, gave rise to digital earths and generated both lively interest in and a disdain for traditional geography found in the neogeography entrepreneurs who developed products like Keyhole, which became Google Earth (Crampton 2009).

Another finding relates to data accuracy. Regardless of data quality, geovisualization gives the impression of greater data accuracy, digital earths chiefly because of the ability to zoom into the particular and navigate the well-defined edges (resolutions) of objects (indeed, we looked for textures to blur edges). Accuracy is not necessarily important; instead, what is essential is confidence in the data. We may not need to know where the trauma occurred especially if it is difficult to obtain higher resolution data than suburban centroid. However, these platforms can misrepresent how much we know.

From a developer’s standpoint, these applications are well-suited for resource-poor environments, which is very important when comparing the cost of digital earth applications to other geovisualization packages. The developer need not have to buy the software nor his/her own hardware. Another potential resource cost concerns the learning curve for developers, which compared to traditional GIS can be considerably less steep. However, if developers need to go beyond simple representations and interoperability then knowledge of more traditional computer programming is vital.

With these applications, development time is reduced, particularly when existing applications can be repurposed. The data can be uploaded quickly and visualized collaboratively. Time may be critical to public health where it concerns the spread of vector borne diseases or a mass trauma event. Conducting health-related work in the cloud has advantages but it exposes data to a proprietary based interface and confidentiality issues. For resource reasons, a psychiatrist may wish to keep all his/her records on Google Docs but governments or businesses can more easily harvest/mine that data. Digital earths may reduce uncertainties in understanding the data but expose one to uncertainties in development.

Conclusions

Many variables affect the visualizations outcomes and many questions arise from our research. Digital earths are rapidly innovating and their future remains difficult to predict but there is a high probability that over time they will better conform to geovisualization principles.

References

Auchincloss, A.H. & Diez Roux, A.V. (2008). A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *American Journal of Epidemiology, 168*(1), 1-8.

Butler, D. (2006). Virtual globes: The web-wide world. *Nature 439*, 776-78.

Chang, Y., Parrales, M., Jimenez, J., Sobieszczyk, M., Hammer, S., Copenhaver, D. & Kulkarni, R. (2009). Combining Google Earth and GIS mapping technologies in a dengue surveillance system for developing countries. *International Journal of Health Geographics, 8*(1).

Cinnamon, J., & Schuurman, N. (2010). Injury surveillance in low-resource settings using Geospatial and Social Web technologies. *International Journal of Health Geographics, 9*(25), 1-14.

Cartwright, W., Crampton, J., Gartner, G., Miller, S., Mitchell, K., Siekierska, E. & Wood, J. (2001). Geospatial Information Visualization User Interface Issues. Cartography and Geographic Information Science 28(1), 45-60.

Crampton J. W. (2009). Cartography: maps 2.0. *Progress in Human Geography* 33(1): 91-100.

Dykes, J. & Wood, J. (2007). Visualization with Google Earth and GIS. http://www.gicentre.org/infoVis/ - Accessed 1/21/2011

Janies, D., Hill, A. W., Guralnick, R., Habib, F., Waltari, E. & Wheeler, W.C. (2007). Genomic Analysis and Geographic Visualization of the Spread of Avian Influenza (H5N1). Oxford Journals, Life Sciences, Systematic Biology 56(2), 321-29.

Le, Quinh. (2005). Issues on health data collection. *Department of Rural Health.* University of Tasmania, Australia.

MacEachren, A.M. & Taylor, D.R.F. (1994). *Visualization in modern cartography*. Oxford: Pergamon Press.

Robinson, A. (2007). A design framework for exploratory geovisualization in epidemiology. *Information Visualization 6*, 197-214.

Sack, D., Brooks, V., Behan, M., Cravioto, A., Kennedy, A., IJsselmuiden, C. & Sewankambo, N. (2009). Improving international research contracting. *Bulletin World Health Organization, 87(*487), 487-88.

Scharl, A. and K. Tochtermann (Eds). 2007. The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society. London: Springer. Information and Knowledge Processing Series.

Schultz, R., Kerski, J. & Patterson, T. (2008). The use of virtual globes as a spatial teaching tool with suggestions for metadata standards. *Journal of Geography 107*(1), 27-34.

Sheppard, S.R.J. & Cizek, P. (2009). The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualisation. *Journal of Environmental Management 90*(6) 2102-117.

Sui, D. Z. & Holt, J.B. (2008). Visualizing and Analysing Public-health data using value-by-area cartograms: toward a new synthetic framework. *Cartographica 43*(1), 3-20.

Zhang, J., Shi, H., & Zhang, Y. (2009). "Self-Organizing Map Methodology and Google Maps Services for Geographical Epidemiology Mapping." *Digital Image Computing: Techniques and Applications*, 229-235.