[CV-05-036] Geovisual Analytics

Abstract

Geovisual analytics refers to the science of analytical reasoning with spatial information as facilitated by interactive visual interfaces. It is distinguished by its focus on novel approaches to analysis rather than novel approaches to visualization or computational methods alone. As a result, geovisual analytics is usually grounded in real-world problem solving contexts. Research in geovisual analytics may focus on the development of new computational approaches to identify or predict patterns, new visual interfaces to geographic data, or new insights into the cognitive and perceptual processes that users apply to solve complex analytical problems. Systems for geovisual analytics typically feature a high-degree of user-driven interactivity and multiple visual representation types for spatial data. Geovisual analytics tools have been developed for a variety of problem scenarios, such as crisis management and disease epidemiology. Looking ahead, the emergence of new spatial data sources and display formats is expected to spur an expanding set of research and application needs for the foreseeable future.

Keywords: analytical reasoning, cartography, geovisualization, spatial data mining, visual analytics

Author & citation

Robinson, A. (2017). Geovisual Analytics. The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2017 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2017.3.6.

Explanation

1. Definitions
2. Introduction
3. Analytical Reasoning
4. Interactive Geographic Visualization
5. Computational Methods
6. Technical Challenges
7. User-Centered Design in Geovisual Analytics
8. Application Contexts

1. Definitions

Geovisual analytics: the science of analytical reasoning with spatial information as facilitated by interactive visual interfaces.

Analytical reasoning: the process of examining information in order to find patterns within that information.

Coordinated-view visualization: dynamic visual interfaces in which multiple data representations are interactively linked to support operations such as cross-filtering,

highlighting, and selection.

Geovisualization (Geographic visualization): interactive visual interfaces for geographic information.

User-centered design: the iterative process of designing a system based on how it will be understood and used by a human user.

Usability: the overall ease of use associated with a given system.

Utility: the ability for a system to solve specific problems.

Needs assessment: the process of identifying and characterizing user needs as inputs to the design of a system.

2. Introduction

Geovisual analytics is the science of analytical reasoning with spatial information as facilitated by interactive visual interfaces (G. Andrienko et al., 2007). This domain emerged in the late-2000s following the development of a broader field called visual analytics (Thomas & Cook, 2005). Visual analytics itself emerged as a separable discipline from the field of Information Visualization, becoming distinct via its focus on the science of analytical reasoning with visual interfaces, rather than primarily emphasizing the visual representation of data as is common in Information Visualization. Put simply, visual analytics places emphasis on supporting analytical processes, not just the development of new visual methods for representing data. Therefore, in the context of geographic information science, geovisual analytics is an emerging research domain that advances the state of the art in analytical reasoning with visual interfaces for spatial data sources, and it builds upon a great deal of earlier work in Geovisualization (see **Geovisualization,** forthcoming) which focused on new visual methods for representing and interacting with spatial data.

Geovisual analytics is a focus today for a significant amount of geographic research on cartography, computational methods, interface design, and cognitive science. The field blends each of these areas together in the pursuit of new interactive mapping systems that allow users to detect patterns and predict future outcomes using spatial data. A key motivating force on this field is its focus on supporting analytical work that connects to complex human and environmental problems. As a result, geovisual analytics approaches tend to embrace the fact that datasets are messy and uncertain (see **Mapping Uncertainty**), that problems may not have singular outcomes, and that different user groups may have radically different requirements for meeting their analytical goals. Geovisual analytics today is also heavily influenced by the challenges posed by volume, velocity, variety, veracity, and other aspects associated with Big Data (see **Big Data Visualization**). A considerable amount of attention in geovisual analytics centers on the challenges associated with analyzing spatio-temporal problems (see **Space-Time Analytics & Modeling**) (Gennady Andrienko et al., 2010).

3. Analytical Reasoning

A key point of differentiation for geovisual analytics is its focus on **analytical reasoning**, the process of examining information in order to find patterns within that information.

Rather than centering primarily on the visualization itself, geovisual analytics calls for scientific progress on supporting people to analyze and make sense of geographic information using visual interfaces. Novel approaches are therefore needed not only for visual interfaces to geographic information, but also for the underlying mechanisms for human reasoning and sensemaking with geographic information. Support for analytical reasoning in geovisual analytics often comes in the form of computational methods that are used to detect patterns and/or predict future outcomes. It can also be evident in interfaces to geographic information that are tailored to support the analytical processes and workflows that analysts use to conduct their work. Research in geovisual analytics therefore includes both ends of the analytical spectrum, ranging from the development of new quantitative methods for pattern detection (Shekhar, Evans, Kang, & Mohan, 2011)to more qualitative studies of analyst sensemaking processes (Pirolli & Card, 2005). Ideally, progress in geovisual analytics comes in the form of improved methods for solving real-world problems, rather than the development of new solutions in search of valid problems.

4. Interactive Geographic Visualization

The visualizations present in geovisual analytics tools can include thematic maps, scatter plots, parallel coordinate plots, timelines, space-time cubes, and a wide range of other representation techniques (Figure 1). Interactivity is usually supported by allowing changes in one view to be broadcast immediately to other available views. Such systems are described generally as **coordinated-view visualizations**. A major stream of research in geovisual analytics focuses on the development and evaluation of new methods for visually representing and interacting with geographic information, and today these efforts include phone, tablet, and large-screen display environments (see **Mobile Mapping & Responsive Design,** forthcoming) in addition to traditional computers.

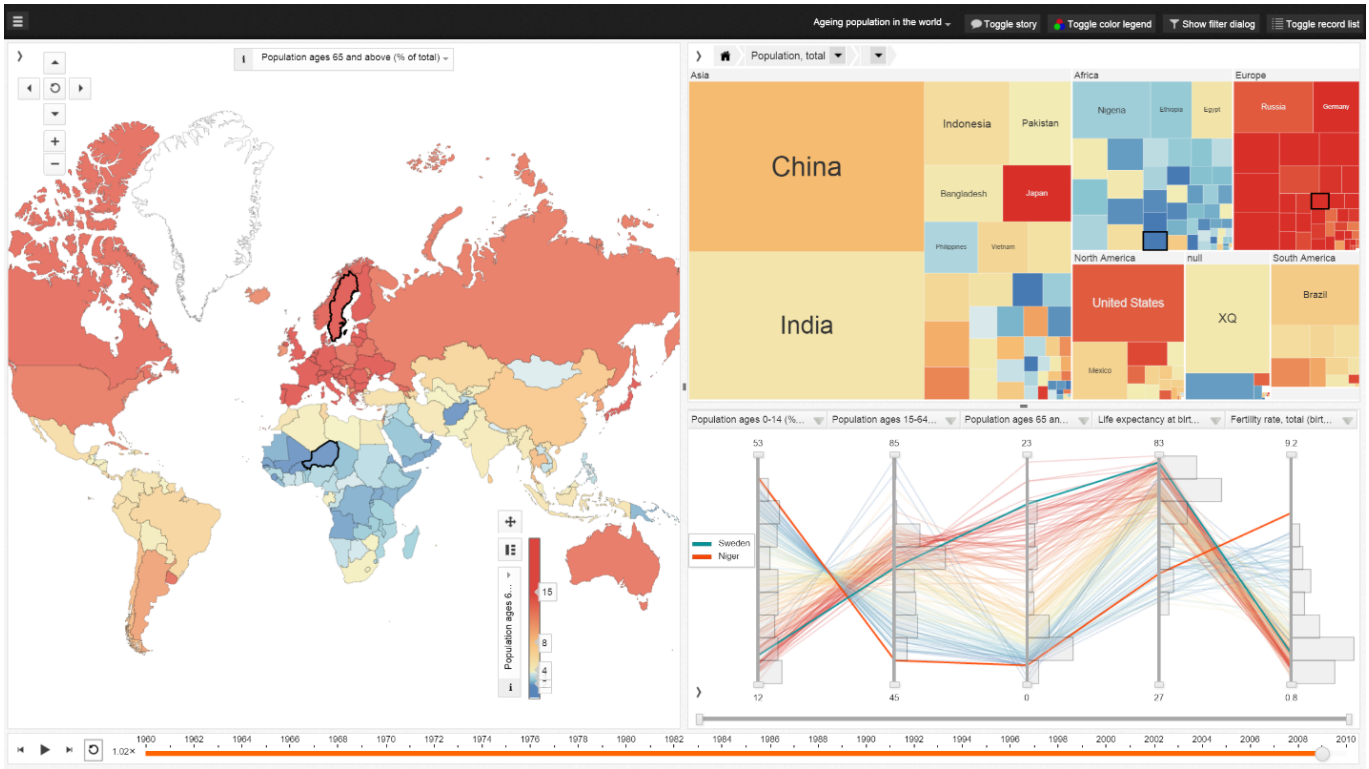


Figure 1: The GAV toolkit is a geovisual analytics system that features a wide range of visualization components that can be used in a coordinated manner to explore multi-variate

spatiotemporal data. In this example, global population trends by country are compared using a choropleth map, treemap, and parallel coordinate plot. Image courtesy of the National Center for Visual Analytics (NCVA) at Linköping University.

Systems for geovisual analytics typically support a high degree of interactivity (Roth, 2013), and this interactivity is directed by the end-user (see **User Interface & User Experience**). Interface controls are normally provided for both the representation methods that are applied to geographic data as well as computational techniques that may be used to extract patterns from complex datasets. A focus on interactivity in geovisual analytics is intended to support iterative engagement by users to evaluate and refine what they are seeing, and therefore enable deductive, inductive, and abductive forms of reasoning (Gahegan, 2005). Interactivity was a central motivating factor behind the rise of **geovisualization** as a domain, the most direct predecessor for what today is known as geovisual analytics. Early progress on geovisualization featured the development of novel methods of map interactivity that allowed users to dynamically modify the visual display of geographic information. Today, in geovisual analytics, interactivity can include the ability to modify visual representations as well as underlying computational methods and tools for supporting storytelling (Van Ho, Lundblad, Astrom, & Jern, 2012).

5. Computational Methods

A wide range of techniques are used in geovisual analytics systems for identifying patterns and clusters in geospatial data (see **Data Mining Approaches**, **Exploratory Data Analysis**, and **Computational Movement Analysis** for examples, although all topics in the **Analytics and Modeling** knowledge area are relevant). In addition, some computational methods in geovisual analytics are focused on predicting future outcomes based on available spatial data. Furthermore, it is now also possible for analysts to leverage pattern matching tools to direct geovisual analytics systems to look for examples based on a previously known target pattern. Each of these approaches relies on the development of spatial analysis methods that can be incorporated into an interactive visualization environment. A variety of scientific and technological problems emerge from this necessity. For example, it is non-trivial to ensure that computational methods process data quickly enough to support real-time interaction and manipulation of analytical parameters. It is also challenging to design interfaces to these methods that are understandable and useful for end-users.

Computational methods in geovisual analytics include cluster detection, geographically-weighted regression, self-organizing maps, named-entity recognition, and many other approaches. The core goal for the application of these methods remains focused on supporting analytical reasoning with spatial information, so the choice of computational methods in geovisual analytics should be directed by analytical needs rather than proposing a solution in search of a problem. It is not uncommon for geovisual analytics platforms to integrate a variety of computational methods in order to tackle different aspects of a given problem domain.

6. Technical Challenges

The development of geovisual analytics systems has moved through several phases of technological development. Early examples were designed for desktop PC environments.

Next, the first wave of web-delivered visualizations emerged, commonly written using Java or ActionScript programming languages. Today, developers have a broad range of options from which to choose. For example, one can use dedicated application program interfaces (APIs) to craft and deliver custom applications based on open source and proprietary GIS platforms. It is also common to leverage cloud-based services to offload computational steps and data storage onto scalable digital infrastructure. A popular choice for designing web-delivered visualizations of all types today is the Data Driven Documents (D3) library (see **Web Mapping**, forthcoming), which is frequently coupled with spatial database frameworks such as PostGIS and statistical analysis libraries like R in order to formulate a fully-functional geovisual analytics environment. Commercial tools such as Tableau and Insights for ArcGIS provide unified frameworks that handle interface, database, and computational needs in a single integrated platform.

To aid the development of new Visual Analytics tools, including geovisual analytics systems, the Human-Computer Interaction Laboratory (HCIL) at the University of Maryland hosts the Visual Analytics Benchmark Repository, which offers dozens of synthetic datasets that feature realistic problems and have known solutions. Because these datasets have patterns and discoveries embedded within them that should be discoverable, they provide excellent fodder for the design of new systems. Furthermore, because they are synthetic datasets, they allow designers and developers to experiment with realistic data for problems which may otherwise involve the use of sensitive or private information. For example, one benchmark dataset simulates patient medical records and a disease outbreak.

7. User-Centered Design in Geovisual Analytics

A central concern in geovisual analytics is to ensure that tools reflect **user-centered design** principles and are iteratively designed with end-user considerations at the core to support high degrees of utility and usability (see **Usability Engineering,** forthcoming). An important distinction is made between the usability and utility of geovisual analytics systems. **Usability** refers to the general ease of use associated with a given system, while **utility** refers to a tool's ability to solve certain problems. It is possible for tools to have a high degree of utility and low degree of usability, and vice-versa. To support analytical reasoning, we must understand the tasks that users wish to accomplish and design interfaces which support that work with ease and efficiency. Research in user-centered design of geovisual analytics systems aims to characterize the extent to which these tools meet user expectations and to draw comparisons with existing methods for completing the same tasks. User-centered approaches to geovisual analytics are iterative in nature, beginning with **needs assessment** to define key user requirements, and continuing with prototype evaluation in cycles until a final design is reached that satisfies user needs and works effectively (A.C. Robinson, Chen, Lengerich, Meyer, & MacEachren, 2005).

User-centered design in geovisual analytics also focuses on how people perceive information in visualizations and how they make sense of their discoveries. For example, perceptual studies in geovisual analytics have involved the use of experimental approaches using eye-tracking to study how people attend to geographic representations. Characterizing the ways in which users assemble and add meaning to their analytical results to tell stories is one example of cognitive research in geovisual analytics. Many other types of user-focused studies are possible with the ultimate goal of providing design guidelines to improve the utility and usability of geovisual analytics systems.

8. Application Contexts

Prior work in geovisual analytics has focused on a wide range of application domains, including crisis management (MacEachren et al., 2011), epidemiology (Maciejewski et al., 2011), political event analysis (Figure 2), mobility and transportation research (Figure 3) (G. Andrienko, Andrienko, Fuchs, & Wood, 2016), and demographic analysis (Slingsby, Dykes, & Wood, 2011). For example, MacEachren et al. (2011) implemented a geovisual analytics system for characterizing the geographic mentions in social media messages compared to the locations from which those messages physically originated. Using a combination of natural language processing and geocoding tools, the places mentioned in text were extracted and located – automating a process that would have previously been done manually by an analyst. Using these enhanced data sources, MacEachren et al. (2011) developed a visual interface that helps users see the spatial footprints of social media conversations about disasters and disaster relief efforts. This project was later extended to support analyzing other topics, such as elections (Figure 4).

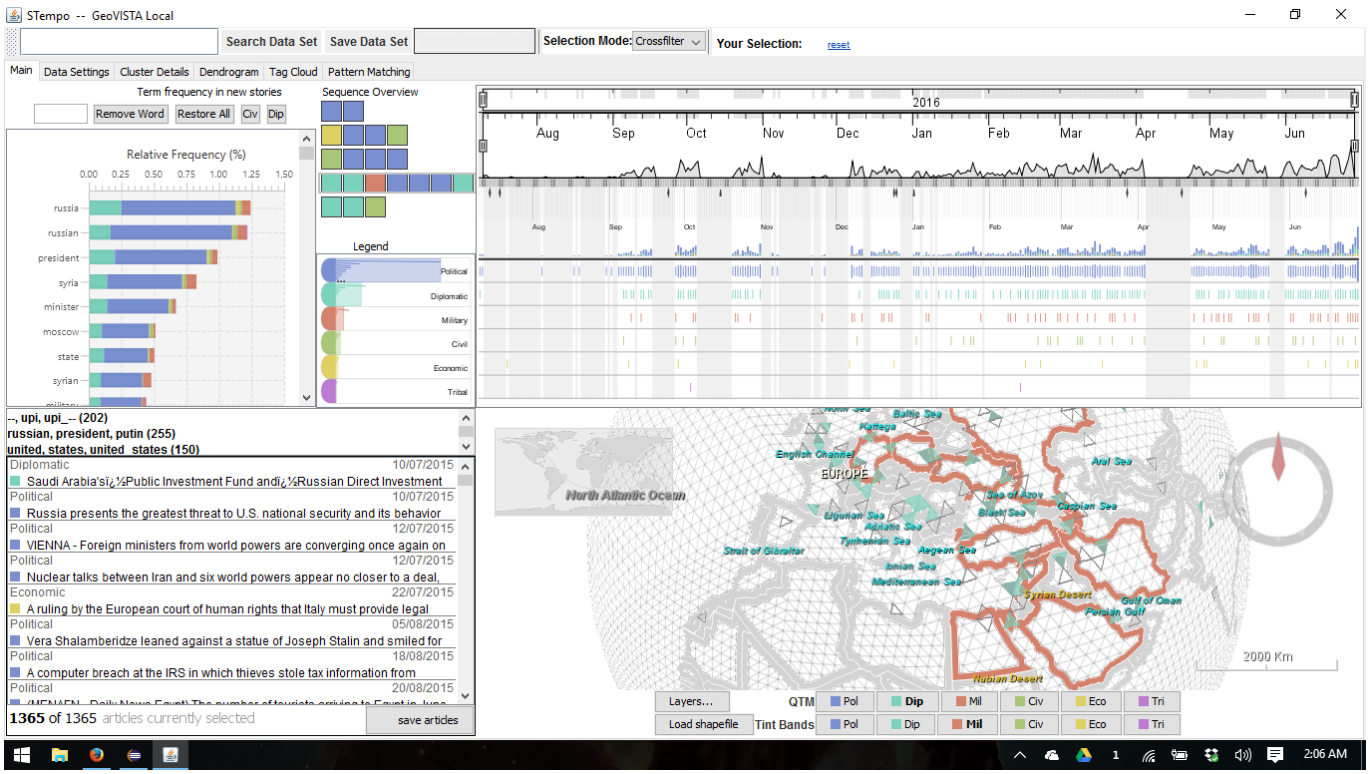


Figure 2: The STempo toolkit is a geovisual analytics system designed to support pattern discovery and analysis in spatio-temporal events collected from news articles (Anthony C. Robinson, Peuquet, Pezanowski, Hardisty, & Swedberg, 2016). This figure shows multiple coordinated views that highlight key terms mentioned in events extracted from news articles, their temporal and spatial distributions, and the results of a modified T-pattern analysis algorithm which suggests potentially interesting combinations of events that could warrant further analyst investigation. These views highlight the spatial, temporal, and attribute footprints associated with newsworthy events in and around Syria in 2015-2016.

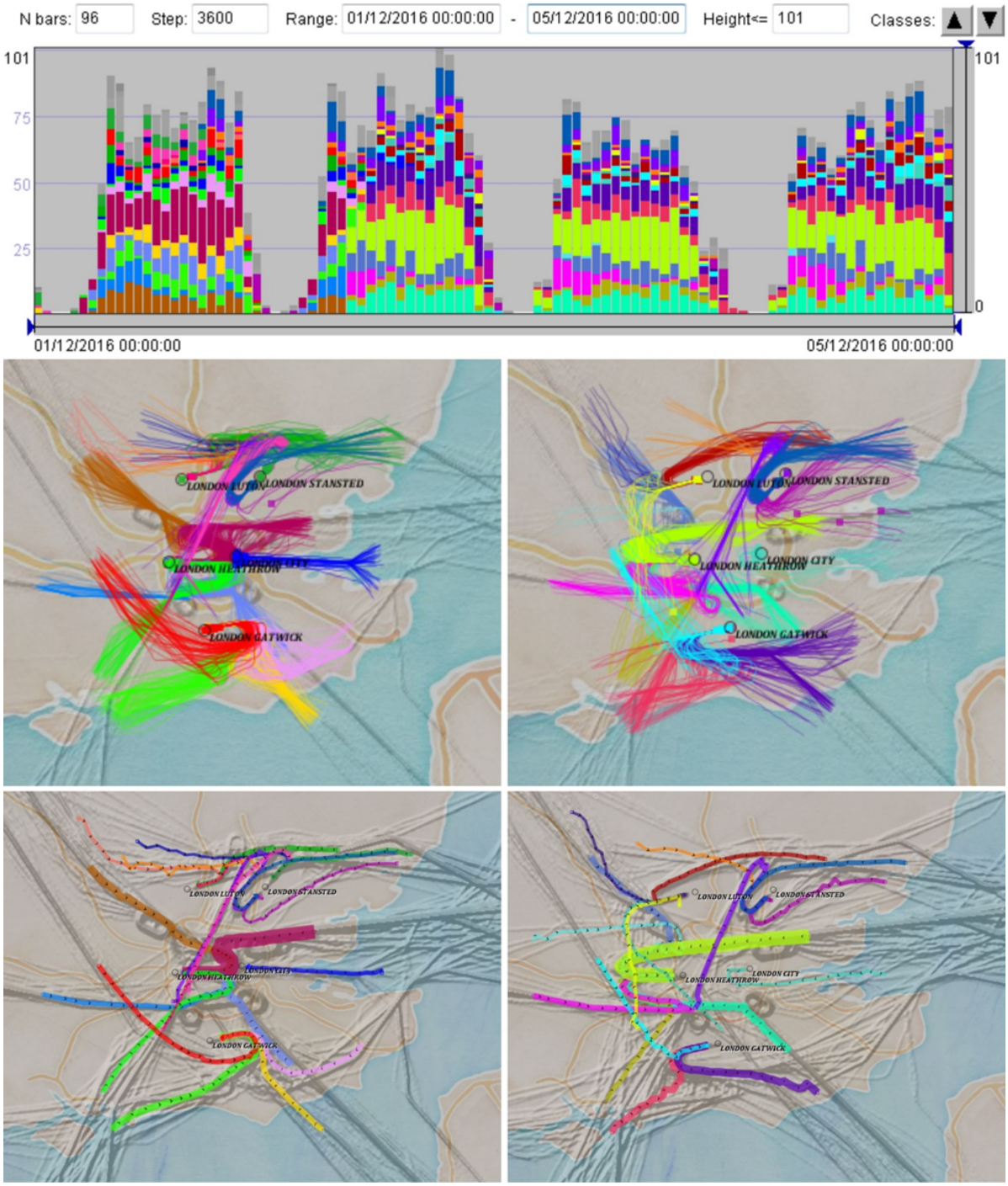


Figure 3. In this example of geovisual analytics for transportation analysis (G. Andrienko, Andrienko, Fuchs, & Garcia, 2017), the bars in a temporal histogram at the top of the image show the number of flight arrivals in hourly intervals for airports in the United Kingdom. The histogram bar segments are colored based on route-based clusters that those flights belong to. Day 1 appears to have a quite different mix compared to days 2-4. At the middle of the image, final segments of each flight's trajectory are colored according to their route-based cluster. On the bottom of the image, days 1 and 3 are displayed using the central trajectory of each cluster to generalize patterns from the overall dataset.

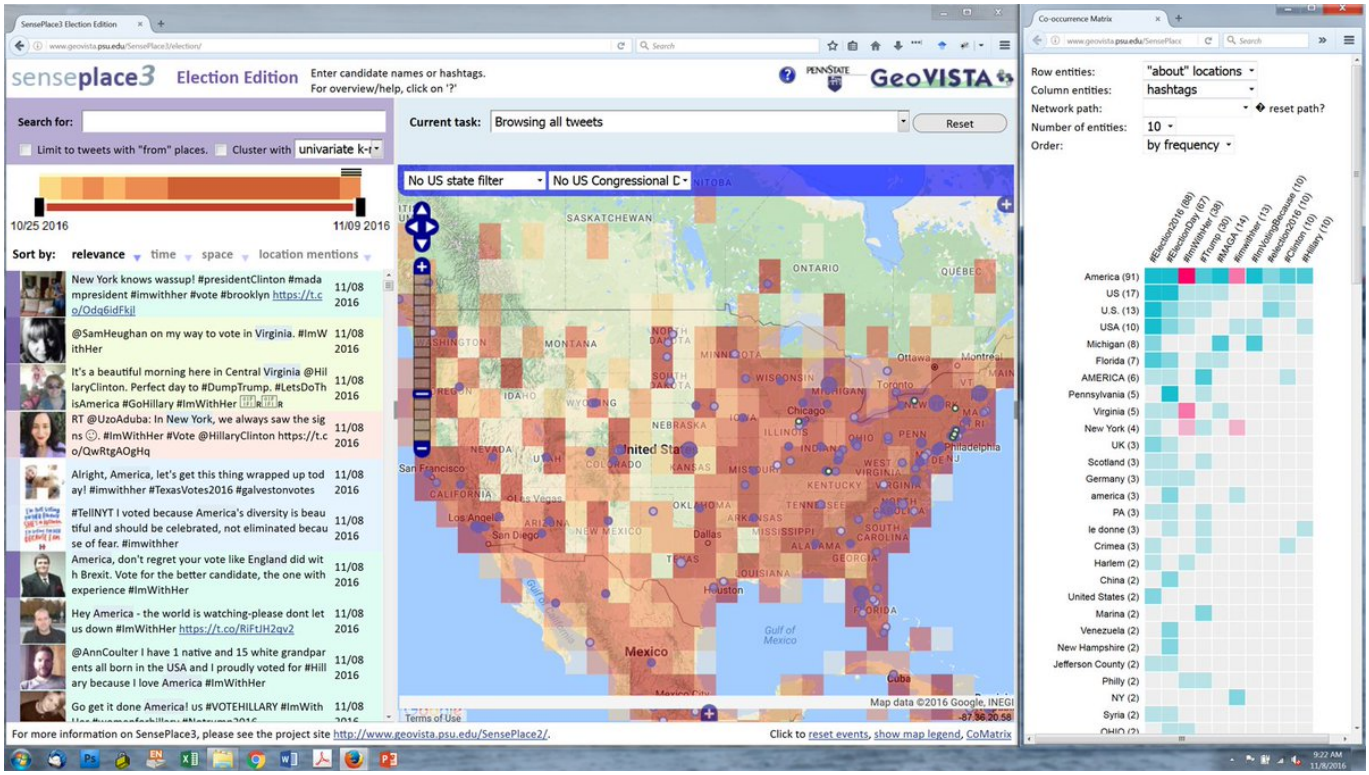


Figure 4. The SensePlace3 geovisual analytics system allows users to explore the place references included in Twitter messages, including the locations messages are sent from as well as those that are referenced in the message content itself. In this example, tweets about the 2016 United States presidential election are visualized from multiple perspectives. A timeline and list view on the left of the display shows recent activity, while the map at the center of the display shows the density of location mentions for election-related tweets. At right, a matrix view allows users to compare relationships between locations mentioned in tweets to the use of specific hashtags.

Future work in this sub-field of GIScience can be expected to build on these existing examples and expand into new application contexts. New approaches to geovisual analytics will be needed in order to leverage the rise of new forms of big data coming from volunteered sources, streaming multimedia, sensors in the Internet of Things, and unmanned terrestrial and airborne platforms. Users of geovisual analytics systems will also expect to engage with these tools on mobile devices, large screens with teams of collaborators, and in virtual environments. Designing new analytical tools and workflows that can adapt to these new spatial data sources and display requirements will provide new research opportunities for the foreseeable future.

References

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., Jern, M., Kraak, M.-J., Schumann, H., & Tominski, C. (2010). Space, time and visual analytics. International Journal of Geographical Information Science, 24(10), 1577-1600.