

Article

Geovisual Analytics Approach to Exploring Public Political Discourse on Twitter

Jonathan K. Nelson ^{1,*}, Sterling Quinn ¹, Brian Swedberg ¹, Wanghuan Chu ²
and Alan M. MacEachren ¹

¹ GeoVISTA Center, Department of Geography, Penn State University, State College, PA 16801, USA; E-Mails: quinn@psu.edu (S.Q.); bws180@psu.edu (B.S.); maceachren@psu.edu (A.M.M.)

² Department of Statistics, Penn State University, State College, PA 16801, USA;
E-Mail: wxc228@psu.edu

* Author to whom correspondence should be addressed; E-Mail: jkn128@psu.edu;
Tel.: +1-505-604-5469.

Academic Editors: Julia Siemer and Wolfgang Kainz

Received: 15 December 2014 / Accepted: 16 February 2015 / Published: 5 March 2015

Abstract: We introduce spatial patterns of Tweets visualization (SPoTvis), a web-based geovisual analytics tool for exploring messages on Twitter (or “tweets”) collected about political discourse, and illustrate the potential of the approach with a case study focused on a set of linked political events in the United States. In October 2013, the U.S. Congressional debate over the allocation of funds to the Patient Protection and Affordable Care Act (commonly known as the ACA or “Obamacare”) culminated in a 16-day government shutdown. Meanwhile the online health insurance marketplace related to the ACA was making a public debut hampered by performance and functionality problems. Messages on Twitter during this time period included sharply divided opinions about these events, with many people angry about the shutdown and others supporting the delay of the ACA implementation. SPoTvis supports the analysis of these events using an interactive map connected dynamically to a term polarity plot; through the SPoTvis interface, users can compare the dominant subthemes of Tweets in any two states or congressional districts. Demographic attributes and political information on the display, coupled with functionality to show (dis)similar features, enrich users’ understandings of the units being compared. Relationships among places, politics and discourse on Twitter are quantified using statistical analyses and explored visually using SPoTvis. A two-part user study evaluates SPoTvis’

ability to enable insight discovery, as well as the tool's design, functionality and applicability to other contexts.

Keywords: geovisual analytics; cartographic display; multiscale visualization; politics; discourse analysis; Twitter; big data; user evaluation

1. Introduction and Political Context

Geovisual analytics has roots in visual analytics, a decade-old interdisciplinary field that focuses on developing, applying and understanding the application of methods and tools that provide visual-computational support for analytical reasoning [1]. Geovisual analytics integrates approaches from cartography and GIScience more broadly with those of visual analytics; it has been described as focusing “...on visual interfaces to analytical/computational methods that support reasoning with/about geo-information—to enable insights about something for which place matters” [2]. Maps are central to geovisual analytics methods, and the prototypical geovisual analytics application uses multiple, dynamically linked views to enable users to explore geographical variation in phenomena of interest. This approach is adopted here in the development of a geovisual analytics tool, the spatial patterns of Tweets visualization (SPoTvis) application. SPoTvis is designed to support the analysis of public political discourse in which difference of opinion is likely and that difference is grounded in place. Here, we present the approach, implementation and the capabilities of SPoTvis through a case study analysis of a divisive political situation in the United States, with a focus on events in fall, 2013, and their reflection in Twitter discourse at the time.

In 2010, the U.S. Congress passed the Patient Protection and Affordable Care Act (commonly known as ACA or “Obamacare”), significantly reforming the country’s health insurance regulations and marketplace. The votes occurred mostly along party lines, with Democrats generally favoring the bill and Republicans opposing it. Debate over funding for the act continued between the law’s passage and its implementation. On 1 October 2013, the U.S. federal government shut down due to disagreement about whether ACA funding would be included in its general fiscal appropriations for 2014. On the same day, the HealthCare.gov online insurance marketplace mandated by the ACA made an ill-fated debut, with long wait times and crashes plaguing the site and impeding signups. The shutdown was eventually ended by congressional compromise 16 days later, while the website was gradually reformed over a period of several months.

In the days prior to, during and after the shutdown, the Twitter conversation surrounding the events seemed just as divided as the politicians’ discourse. The hashtags #shutdown and #Obamacare were often seen trending during this period, with acrimonious tweets directed at the instigators of the shutdown, as well as the proponents of the ACA; however, it was difficult to tell where these veins of messages were originating and whether they were always coming from the same people. As we scanned the online conversation during this period, two main research goals emerged. First, we wanted to determine if other subthemes could be identified beyond the general hashtags of #shutdown and #Obamacare. Was there any conversation from furloughed government workers about household financial struggles? Were people talking about closures of national parks and monuments? Were people

mentioning troubles signing up for insurance on HealthCare.gov? We hoped to get beyond the noise to find the most poignant subjects of concern for U.S. residents.

Second, we wanted to uncover spatial patterns in the usage of these subthemes. Did the patterns of support and opposition to the shutdown follow known geographic regions of Republican and Democratic party dominance? Did the public discourse in a congressional district match the political position of the district's congressional representative? What subthemes were of most concern in different areas of the country? What regions were talking about the most similar things?

The remainder of this paper describes the development, application and assessment of SPoTvis as a web-based geovisual analytics tool that helps users to uncover some answers to the above questions. The tool offers a term polarity plot coupled with a pair of interactive maps that allow users to compare Twitter subthemes between any two states or congressional districts (Figure 1). The themes are weighted and placed on the term polarity plot's horizontal axis that represents a continuum of interest between the two districts. Terms near the left and right edges of the axis tend to only appear in one of the two political units, while terms near the center of the axis experience more balanced usage patterns between the two units. Demographic attributes and indicators of partisanship provide additional contextual information that enriches users' understandings of how the conversation follows political leanings. Finally, an option to "see similar districts" helps to determine whether regional patterns exist in the discourse. As users explore the map and exercise these functions, they can increasingly make sense of the way varying opinions and values appear in a social media conversation across space.

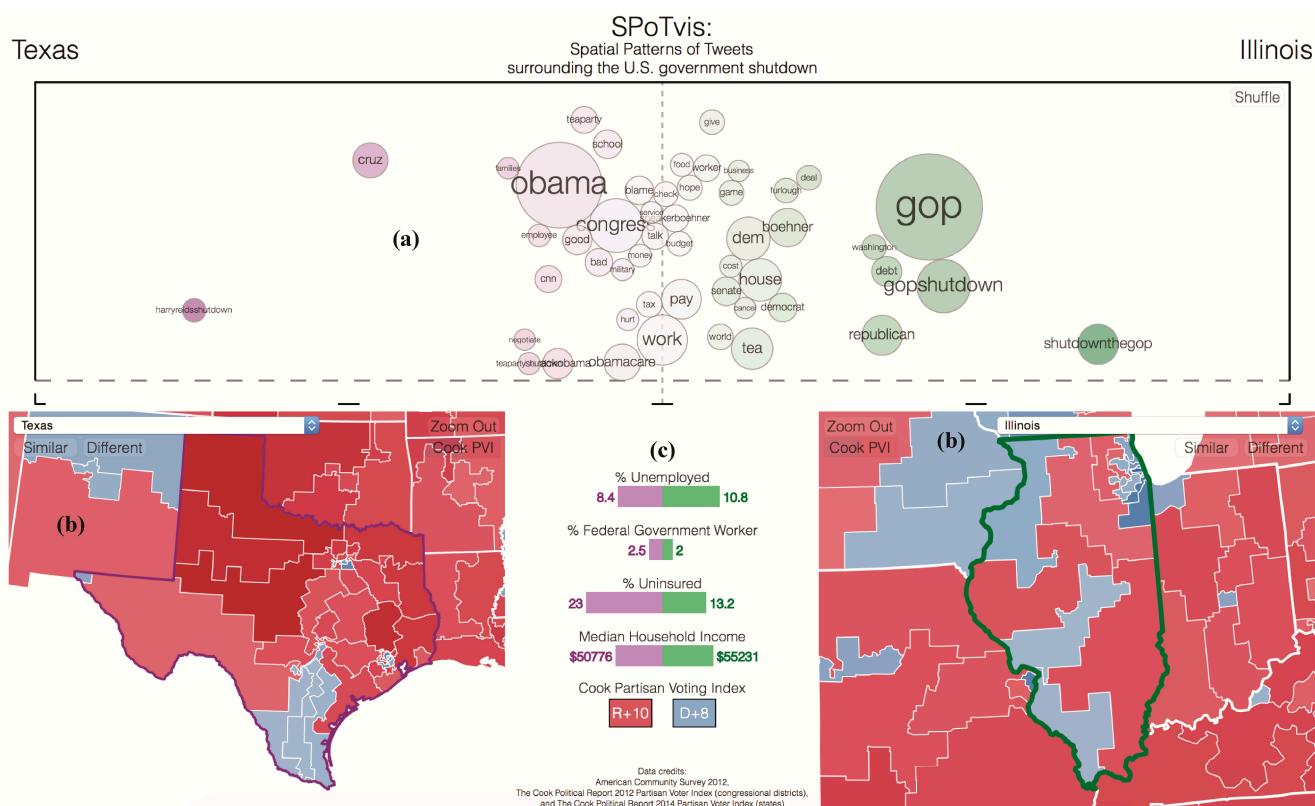


Figure 1. Screenshot of SPoTvis [3]. **(a)** Term polarity plot; **(b)** map views; **(c)** demographic panel. PVI, Partisan Voter Index.

In the following sections, we first justify the value in exploring Twitter data using visual analytics applications. We then outline our data collection process and present findings from a statistical analysis on the relationships among places, politics and discourse on Twitter in the context of the 2013 U.S. Government shutdown. Next, we discuss the rationale behind the design of SPoTvis, how the tool was implemented and considerations for its performance and usability. We provide brief results on spatio-political patterns found using SPoTvis. We then evaluate SPoTvis' ability to enable insight discovery, as well as the tool's design, functionality and applicability to other contexts using a two-part user study. Lastly, we conclude with a summary of the findings and exciting future directions for SPoTvis development.

2. Justification for Visual Analytics and Twitter Use

Journalists and scientists are using social media data increasingly to uncover newsworthy stories or to better understand what people talk about and how they share information. For example, researchers have shown that the public can generate and spread information during natural disasters and other emergencies using social media outlets, such as Twitter [4]. It is also possible to gauge disease activity [5] and public sentiment on climate change [6]. While Twitter provides an endless stream of conversations for potential study, making sense of and drawing conclusions from such large data sources is challenging. Incorporating the geographic component of tweets, or where people tweet from, further complicates this process. As noted above, visual analytics provides tools and frameworks for addressing these challenges [1,7].

In the virtual Twitterverse, political conversation follows a unique structure and happens among a special group of Twitter users that is not representative of all users. Smith *et al.* [8] mapped thousands of Twitter topic networks and identified six different kinds of network crowds based on distinct social structures. The authors found that heated political subjects often result in polarized crowds that ignore one another and share different resources. Party affiliation often divides the crowd, and although both crowds are focused on the same topic, little engagement happens between them. Diakopoulos *et al.* [9] explored how political discord tweeted during broadcast events can be used for journalistic inquiry in their tool VoxCivitas. The authors collected tweets during President Obama's 2009 speech at the Copenhagen 15 meeting and extracted keywords and sentiment from them. VoxCivitas allows users to search and filter by prominent keywords tweeted during the speech. The tool displays sentiment on a timeline to gauge whether the Twitter reaction during that time in the speech was positive, negative, controversial or neutral. While SPoTvis does not calculate or report an explicit measure of sentiment, it does display a unique signature of conversation topics exhibited by any two political units, an approach that could complement standard sentiment analysis. SPoTvis also extends beyond the scope of VoxCivitas by considering the spatial component of political tweets, including direct comparison between places, as well as the ability to find both similar and dissimilar places based on term usage.

Other recent research directly related to SPoTvis has used keyword analysis and dynamic visualization approaches to extract political topics and sentiment from Twitter data. Xu *et al.* [10] used Twitter data from the 2012 U.S. presidential election, as well as the Occupy Wall Street movement to demonstrate the agenda-setting role of elite “opinion leaders” in social media conversation. The authors assessed public competition of topics associated with the two political events using topic modeling and

integrated storyline-ThemeRiver [11] visualizations. Tsou *et al.* [12] also used the 2012 presidential election as a case study, but explored how keywords related to political candidates change in response to major events. The authors clearly documented how online conversations about political candidates shift after a major event, such as a hurricane, and how the discussions relate to political candidates. The authors also mapped the proportions of tweets about Barack Obama vs. Mitt Romney before and after such events; they found that the geographic discussion about the candidates does not simply mirror the voting map.

While research on using and mapping social media is undoubtedly growing, leveraging Twitter data remains challenging. Simply displaying tweets on a map leaves much to be desired in terms of cartographic design and the ability to identify patterns [7]. However, visualizing social media topics is important because online conversations are manifested in complex and unpredictable ways. We have yet to find any analysis or tool that compares two spatial units based on their Twitter content, a place in the literature that SPoTvis fills. The combination of interactive maps and visual elements implemented in SPoTvis offers a path toward sensemaking and meaningful visualization and analysis based on large geo-social data.

3. Data Collection and Analysis

The data for this application consists of spatial data representing political boundaries in the United States (at state and congressional district levels), demographic and political data that provide cultural context and tweets about the government shutdown and the ACA.

3.1. Spatial Data and Demographic Attributes

Spatial data in this project consisted of U.S. state boundaries, as well as district boundaries for the U.S. House of Representatives. The U.S. has a bicameral legislature in which members of the Senate are elected on a statewide level and members of the House of Representatives are elected based on a geographic district that can potentially be smaller (but not larger) than a state boundary. The use of the two scales of spatial aggregation allows users to compare the political climate for both senators and house representatives. It also allows users to detect both broad regional differences and more local variation in conversation themes.

To help users contextualize the patterns in tweet content, we collected demographic attributes for states and house districts that may provide insight into how people view the shutdown and Affordable Care Act. The median household income and percentage of unemployed persons (age 16 and over) give a general sense of the economic situation and political concerns in a given district or state. The percent without health insurance may lend insight into how people feel about the Affordable Care Act. Finally, the percentage of employed population that work for the federal government gives an idea of what proportion of the state or district was potentially furloughed during the shutdown. Users can then make inferences about whether the financial and emotional toll imposed by the furloughs affects the attitudes reflected by the tweets.

The latest available Cook Partisan Voter Index (PVI) for each unit was also recorded to help users understand the political leanings of states and districts. This index takes into account the unit's voting behavior in recent presidential elections and assigns the unit a score based on how "Republican" or

“Democrat” the unit is compared with the other units in the United States. For example, the state of Nevada has a Cook PVI of D+2, meaning that it leans slightly Democratic, while the state of Wyoming’s Cook PVI of R+22 means that the state leans heavily Republican. Used frequently by think tanks, pundits, and journalists, the Cook PVI aids users’ understandings of the general political persuasions of people generating tweets in each state and district. Although the segment of the population producing geolocated tweets may not exhibit the same Cook PVI as the entire voting population of the unit, the Cook PVI provides a relative measure of partisanship that is still informative for analysis.

3.2. Tweet Collection

The tweets used in this research were collected continuously during a period of eight weeks between 1 September and 27 October 2013. This period includes much of the congressional ACA funding discussion, the shutdown itself and some of the shutdown aftermath. During this entire period, the ACA was a major topic of focus in both news media and social media. For further details of the tweet collection, see Bodnar and Salathé [13].

To be included in SPoTvis, a tweet needed to contain exact geographic coordinates (in other words, the originator of the message explicitly opted-in to the use of location services). A tweet also needed to contain one of the following keywords: #shutdown, shutdown, #ACA, ACA, Affordable Care Act, #Obamacare, Obamacare and healthcare.gov. Terms such as “website”, “health” and “insurance” were not considered, because they were deemed too general and might introduce themes unrelated to our topics of interest. This dataset resulted in approximately 70,000 tweets within the boundaries of the contiguous United States that could be used for analysis.

3.3. Statistical Analysis of Tweets

Tweets were spatially joined to the vector boundary files mentioned above in order to perform analysis at both the state and house district levels. Of all states, California had the most tweets, with 7169, and Vermont had the fewest, with 90. Because a small percentage of Twitter users enable location services, we acknowledge that our collection represents only a small sample of the tweets generated during the study period. Nevertheless, the histogram in Figure 2 shows that most states had somewhere between 200 and 1600 tweets. The number of tweets per house district ranged from 42 to 807, with the District of Columbia as an outlier with 4005. The histogram in Figure 3 shows that most districts yielded a sample of 100 to 200 tweets.

The textual content of tweets was analyzed using the R statistical software, version 2.15.1, using the “tm” package. We imported the text of the tweets in the corpus data format. The preprocessing of the raw tweets included converting text to lower case, removing all English stop words, punctuation and URLs and stripping extra whitespace. Porter’s stemming algorithm was then applied to convert inflected words to their root forms.

A raw word count was first performed on the tweets to determine the most frequent terms in the dataset. To maintain consistency when building the term polarity plots for the units (described later in this paper), we manually selected 50 of the most common keywords that were meaningful in the context of our research, such as “obama”, “gop” (a nickname for the Republican party), “congress” and “furlough”. The broad term “shutdown” (present in most of the tweets) and obscure words, such as “can”

and “still”, were omitted from this keyword list in favor of terms that were more descriptive of subthemes in the conversation, such as “school”, “worker”, “tax” and the names of politicians.

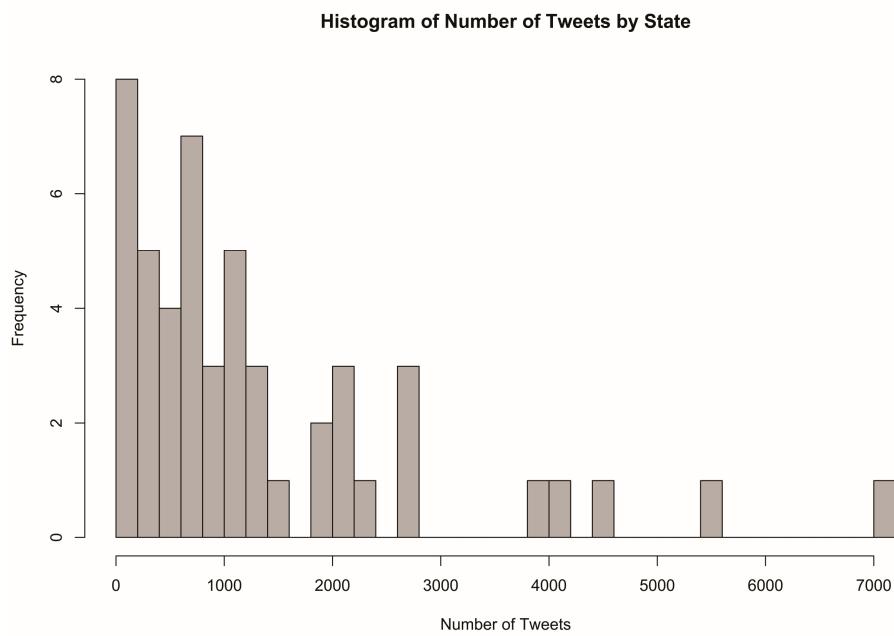


Figure 2. Distribution of tweets by state.

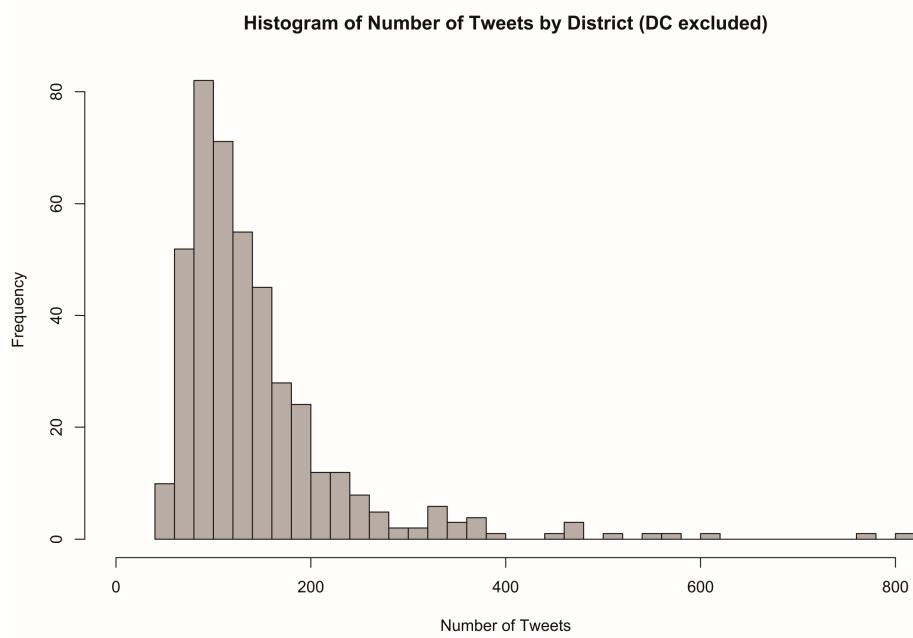


Figure 3. Distribution of tweets by congressional district.

The popularity of a term was measured by the number of tweets containing it. Table 1 shows the top 10 keywords at the national level. For a given unit, in other words, a state or a congregational district, we can obtain a popularity measure for each keyword. Thus, we can construct a vector for the unit of interest:

$$P(\text{unit}_i) = (F_1, \dots, F_N) \quad (1)$$

where $F_k = C_k/n_i$, C_k being the total number of tweets containing the k -th keyword and n_i the total number of tweets in unit i . Figure 4 shows the popularity of 10 keywords for the second and third districts of Alabama (coded as District 102 and 103, respectively). Note that the biggest difference occurs in the words “gop” and “obama”. People in the second district tweet frequently about “obama”, but very little about “gop”. In this particular comparison of districts, no other keywords exhibit such a large variance. People in the third district use the terms “obama” and “gop” with a frequency that is much closer to being equal.

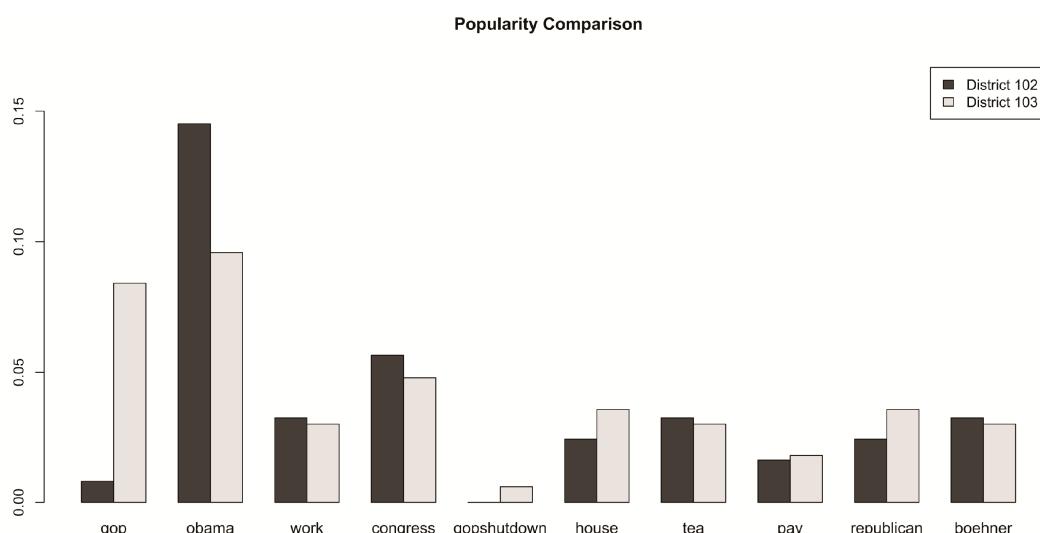


Figure 4. Keyword popularity between the second and third districts of Alabama (coded as District 102 and District 103, respectively).

Table 1. Ten most frequent terms in the Twitter dataset.

Rank	Term	Frequency
1	gop	7699
2	obama	7256
3	work	3877
4	congress	3743
5	gopshutdown	2919
6	house	2734
7	tea	2725
8	pay	2423
9	republican	2405
10	boehner	2367

A unique feature of SPoTvis is that it allows users to explore the set of districts that exhibit similar keywords in their tweets. By previous construction, we have for each unit an $N \times 1$ vector as the keyword popularity measure, where N is the number of keywords under consideration. Thus, the similarity between two districts can be measured by their distance in this $N(=50)$ -dimensional keyword space. While there are many similarity/distance measures in the existing literature, we chose the commonly-used Euclidean distance. For $P(\text{unit}_i)$ and $P(\text{unit}_j)$ defined above, their Euclidean distance is defined as:

$$ED_{ij} = \sqrt{(F_{i1} - F_{j1})^2 + \dots + (F_{iN} - F_{jN})^2} \quad (2)$$

3.4. Examining Links among Places, Politics and Tweets

We examined the above similarity measures to explore whether units with similar keyword usage patterns exhibit any detectable patterns in partisan leanings. We hypothesized that a unit's Cook PVI should be positively correlated with the Cook PVIs of its similar units (here, we mean "similar" in terms of keyword usage patterns), but negatively correlated with its dissimilar units. We chose U.S. states as the units of interest and calculated for each the average Cook PVI of its 10 most similar states (Y_{1i}) and also the average Cook PVI of its 10 most dissimilar states (Y_{2i}). We constructed the following two regression models:

$$Y_{1i} = \alpha + \beta_1 X_i + \epsilon \quad (3)$$

$$Y_{2i} = \alpha + \beta_2 X_i + \epsilon \quad (4)$$

where X_i is the Cook PVI of state i . The corresponding hypotheses are $H_1: \beta_1 > 0$ and $H_1: \beta_2 < 0$.

The estimation result is shown in Table 2. Both of the coefficients are statistically significant, and their signs are as we expected. Note that the magnitudes of these two coefficients, as well as the R-square values, are not very large. This is reasonable, because many other factors could influence the Cook PVI of the similar states. Shown in Figure 5 is the scatter plot of the data with fitted regression lines. The blue and red represent the data of similar and dissimilar states, respectively. From these results, we can conclude that states with similar patterns of keyword usage in their conversations tend to have similar partisan leanings. This result further reinforces the polarized crowd structure found by Smith *et al.* [8] in virtual (rather than spatial) network space.

Table 2. Linear regression estimation result.

	Estimate	p-Value	R-Square
β_1	0.13913	0.00199	0.1895
β_2	-0.10342	0.00706	0.1474

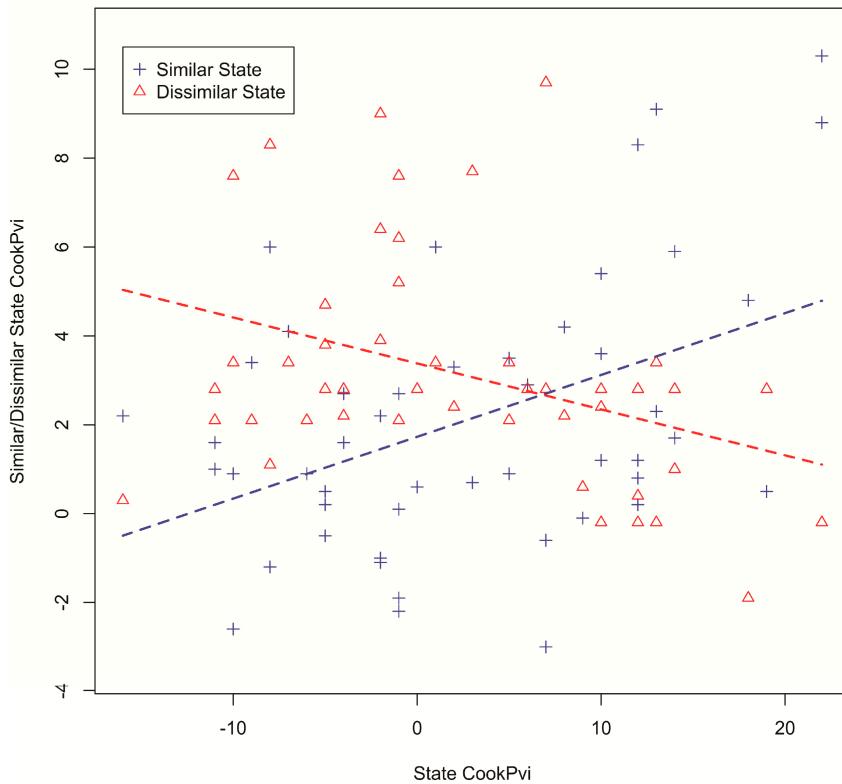


Figure 5. Scatterplot with fixed regression lines comparing a state’s Cook PVI with the average Cook PVIs of its 10 most similar states (blue) and 10 most dissimilar states (red).

4. SPoTvis Design, Functionality and Use

In this section, we discuss the rationale behind the design of SPoTvis. We explain how the tool was implemented and the considerations made for achieving the best performance and usability. Lastly, we present findings based on visual data exploration using SPoTvis.

4.1. Design Rationale

Central to the design of SPoTvis is comparison: “What shutdown-related keywords were being tweeted in my congressional district, as compared to, say, neighboring districts? What about all districts held by a certain political party? Additionally, can I compare my district with statewide trends?” These were the types of questions we strove to visually and interactively address. Additionally, we wanted to supplement this keyword and spatial comparison with demographic information, such as unemployment, income and health insurance enrollment aggregated to the same levels of comparison.

These goals required a duality approach to interface design; one that forces users to input two parameters in order to explore a comparison between two entities, two groups of entities or one entity and a group. Our aim was to capture the analytical and insight-seeking interests of social scientists and news journalists, but also the social visualization interests of society at large. The primary components of the application include a term polarity plot, two spatial map views and a simple graph panel for demographic statistics (refer to Figure 1).

The core component of the interface is the term polarity plot. The plot combines a relatively newer text visualization technique, the word cloud, with proportional symbols, a more traditional method for

visualizing the magnitude of quantitative data. One of the earliest word cloud implementations, a collective mental map of landmarks in Paris, dates back to 1976 [14]. The first usage of proportional symbols dates back far earlier to William Playfair's 1801 statistical graphics that scaled the areas of countries relative to the areas of circles [15].

Independently, these techniques for data visualization remain quite prominent in today's interactive web environment. An early usage report on Many Eyes [16], a public website developed by IBM that allows users to freely upload their data and create interactive visualizations, revealed proportional symbol plots and word clouds to be the most and third most used graphics, respectively [17]. Moreover, users of Wordle [18], a web-based tool for visualizing text, create a new wordle about every ten seconds [19]. Wordles advance the standard word cloud through innovative text placement algorithms and attention to aesthetics, but lack interactivity and semantic word placement.

In visual analytics, word cloud text visualizations have been used to track content evolution in documents through time [20] and to compliment more advanced text visualizations, such as TextFlow, in depicting how topics evolve in large text collections over time [21]. Word cloud visualizations have been combined with parallel coordinate plots to reveal regional and linguistic differences between U.S. Circuit Court decisions [22]. However, few interactive visual analytics tools interlink the word cloud and proportional symbol techniques, with a notable exception by Mike Bostock *et al.* [23] of *The New York Times*. In an online article, the authors visualized the comparison between words used by Democrats and Republicans at the 2012 National Convention. Font and bubble size conveyed cumulative word use frequency, and bubble color relayed the proportion of use by political party.

Our term polarity plot extends the work discussed above by integrating the spatial component of text data through bubble placement and dynamic linked map views. The term polarity plot visually encodes keyword frequency for two compared units in a single interactive graphic. For any comparison, the 50 designated shutdown-related keywords appear atop of and are inherently linked to dynamic bubbles. Font size and bubble area are scaled linearly in proportion to the total number of times the word was tweeted within the two compared units.

Horizontal bubble positions are derived based on the relative occurrence between spatially-driven comparisons and placed on a diverging continuum. For example, a bubble within the center of the plot indicates similar use between the two spatial units being compared, whereas a bubble plotted at the far left or right end of the plot indicates a word predominately occurring within only one of the compared units. Politically neutral terms, such as "congress" often appear toward the center, whereas other, more politically charged terms are pulled toward the edges of the plot. An example from the initial map view is "harryreidsshutdown" (a derogatory reference to the Democratic Senate Majority Leader, Harry Reid), which appears far to the side of the Republican-leaning states. It is important to note that bubble placement in the term polarity plot represents similarity in the usage of individual terms between two spatial comparisons; whereas, the show "similar/different" functionality returns spatial entities that are most similar or different in patterns of term usage across all 50 keywords.

Bubble placement in the vertical direction is based on a Gaussian curve to more intuitively fill the plotting area, but portrays no spatial or semantic context. Users can shuffle the vertical placement of bubbles for better clarity in the event that bubbles and text become cluttered and illegible. For further clarity, or to focus on a particular subset of keywords, users can remove uninteresting bubbles from the plot by clicking on them. These demoted bubbles are reduced in size and drop vertically to the bottom

of the plot. Demoting the largest bubble in the plot re-scales the remaining bubbles to foster more visually prominent word comparisons. Re-clicking demoted bubbles reintegrates their size and placement within the plot. To focus on one term, users can hover over the bubble of interest. This highlights the bubble and slightly expands the font size of the word, while also decreasing the visual emphasis of non-highlighted bubbles through reducing their bubble and font opacities.

To visually reinforce the importance of bubble position in relation to keyword frequency comparisons, a diverging color scheme suggested by Harrower and Brewer [24] was implemented. This color scheme further serves to visually link data views (other than partisanship) through the map and statistical graph components of the tool.

SPoTvis' two map views provide a spatial context to the keyword comparisons and allow users to explore the patterns exhibited by different states and congressional districts. To manage the quantity and complexity of information displayed in the map views, we used a multiscale visualization approach to information design. Data and data representation are abstracted at different zoom levels in accordance with Shneiderman's information seeking mantra: "Overview first, zoom and filter, then details-on-demand" [25]. For example, users first click a state, which updates the term polarity plot to a keyword comparison at the state level, zooming the user to the geography of the state. At this zoom level, congressional districts become visible, and the user can choose to click one, again updating the plot. Alternatively, users can filter using drop-down menus, which contain all possible state and congressional district aggregates, as well as a unique option to select all Democratic- or Republican-leaning states (Figure 6).

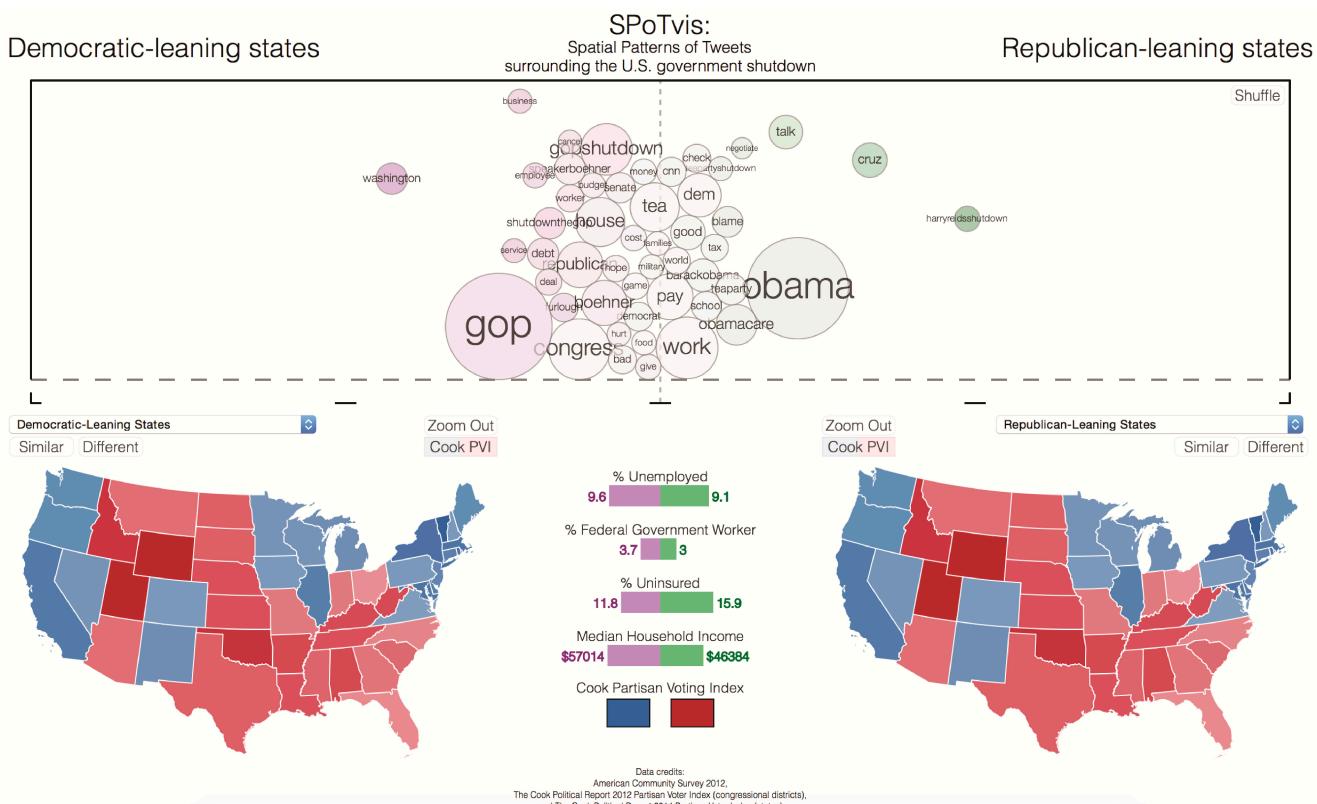


Figure 6. Initial view of SPoTvis comparing Democratic-leaning states with Republican-leaning states.

Additionally, the map views provide options to display the ten most similar or dissimilar states or congressional districts to an entity of interest. Similarities are depicted in the color linked to the left or right comparison views, purple or green, respectively. To alleviate any confusion, dissimilarities are shown in orange, a color intentionally not linked to any other aspect of the interface. If a user clicks to reveal similar or dissimilar geographic entities, links to these areas are displayed. Because some of the smaller congressional districts are hard to discern at the national scale, users can hover over the ten links to flash the centroids of the similar or dissimilar areas and click the links to zoom to these areas. The map views also allow users to toggle on and off Cook Partisan Voting Index (PVI) value maps at both the state and congressional district levels.

Finally, bar charts display the percent of federal government workers, the percent of unemployed, the percent of individuals without health insurance and median household income for any two political units being compared. The charts are aligned horizontally to portray the relative balance between values and to preserve screen real estate. Cook PVI values are also displayed in this area to coincide with the color-coding of their representation in the map views.

4.2. Implementation

SPoTvis was written as a web application in JavaScript, HTML and CSS to make the tool available to any interested, politically-aware Internet user. D3 (Data-Driven Documents) was the primary JavaScript library used for development. D3 binds document object model (DOM) elements to raw data [26]. By doing so, the library operates within the environment of the DOM rather than wrapping the DOM in code in a cumbersome scheme. Furthermore, D3 simplifies element creation, update and removal, which allows for an efficient way to create complex and dynamic visualizations.

Besides handling DOM elements, D3 has many convenience features that are empowering in the development process. In the term polarity plot, a force layout is imposed on the bubbles. This physical simulation uses pseudo-gravity and charge to pull bubbles toward their respective normalized x-values. If conflicts arise, they are handled in an organic way. This process determines bubble position.

TopoJSON, a topology-preserving extension to GeoJSON, was used in combination with D3 to produce the map views. TopoJSON was chosen, because it uses inferred topology to significantly reduce the number of vertices downloaded to the client, thereby shortening the load time. Furthermore, by maintaining topology, we reduce the likelihood of visibly misrepresenting complicated areal units, such as congressional districts.

4.3. Performance and Usability

To keep performance smooth for this prototype tool, all keyword frequencies, similar districts and demographic attributes are pre-calculated and stored in a single JavaScript Object Notation (JSON) file for each state and congressional district. The geometries of the states and districts are sent to the client at the time the application loads. These boundaries are generalized as much as feasible in order to reduce the amount of data transferred. Apart from the initial application load and the occasional retrieval of the state and district JSON files, the application does not need to make any other queries to the server, and no calculations are performed on the fly.

When the term polarity plot undergoes an update, sometimes the bubbles will get “stuck” as they attempt to switch positions on the plot. The bubble color always reflects the exact position that the bubble should occupy on the x-axis, thereby indicating whether the bubble is out of place. The user can drag and drop a stuck bubble to the correct side of the graph, remove irrelevant bubbles to create more space or click the “shuffle” button to assign all of the bubbles a new position on the y-axis, thereby opening some space for movement.

4.4. Data Exploration

Using SPoTvis to explore various states and congressional districts reveals some patterns that indicate how an area's Twitter conversation may be influenced by local current events and political leanings. Comparing just about any combination of states or districts reveals a blame game at play. Units with a highly Republican Cook PVI tend to use the term "obama" more often, and units with a highly Democratic Cook PVI tend to use the term "gop" more often. This pattern is readily visible when comparing polarized states, such as Texas vs. Massachusetts (Figure 7a), but can even be seen when comparing more moderate states, such as Nevada vs. North Carolina, or districts within a state, such as Washington 7 (containing the liberal city of Seattle) with Washington 5 (containing the more conservative city of Spokane and the surrounding area). The latter comparison is shown in Figure 7b.

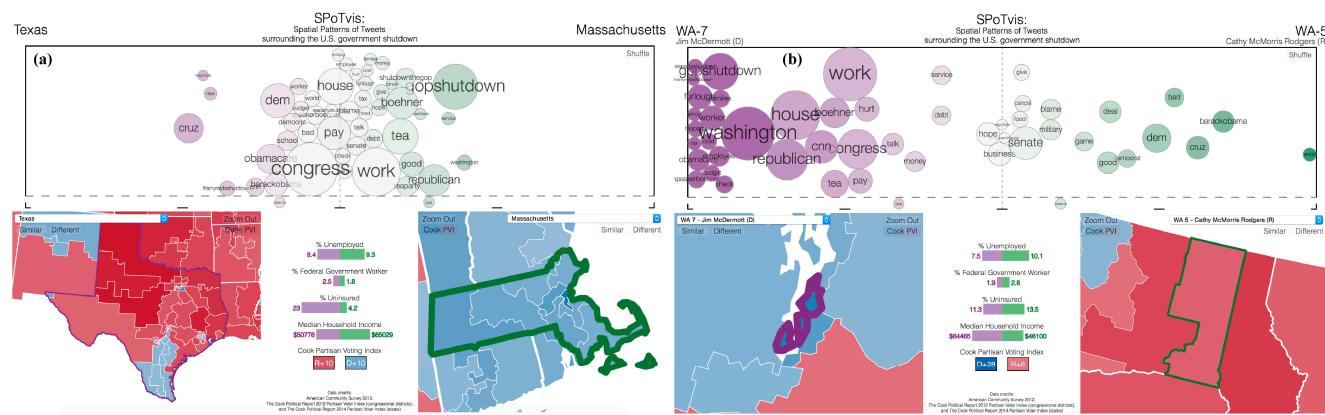


Figure 7. A blame game at play. (a) SPoTvis comparison of Texas and Massachusetts; (b) comparison of Washington 7 and Washington 5.

Local issues also make their way into the term polarity plot. These patterns are more easily seen if the user rescales the plot by removing the dominant terms of “obama” and “gop”. Figure 8 shows the detail of the term polarity plot comparing Maryland (left) with New Jersey (right). These states are geographically close and have a similar Cook PVI; however, the sentiment of the much larger percentage of federal workers in Maryland is captured by the terms that indicate impact, with “work”, “washington”, “furlough”, “employee”, “workers”, “check” and “families” more dominant on the left side of the plot (along with some specific attention to the GOP speaker of the house, Boehner). In New Jersey, the focus is less on outcomes and more on blame, with more instances than in Maryland of terms “tea”, “gopshutdown”, “republican”, “obamacare”, “blame” and “shutdownthegop”, with “school” as the only somewhat prominent term not linked to blame.

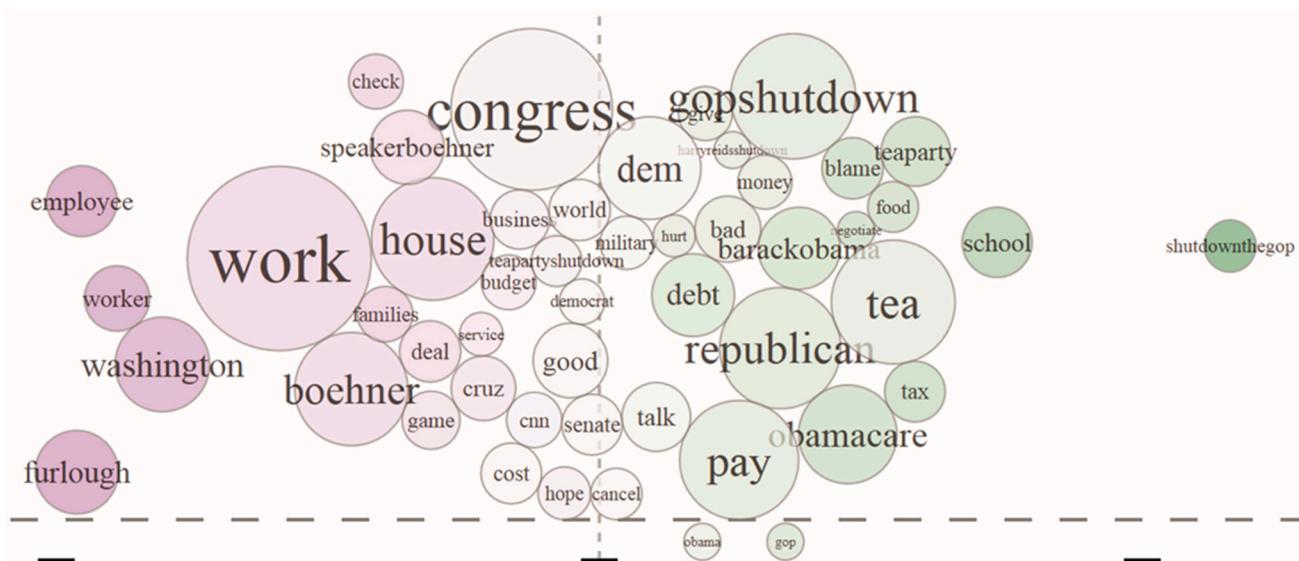


Figure 8. A detailed view of the term polarity plot comparing Maryland (**left**) with New Jersey (**right**) exposes the concerns of the federal workforce in Maryland.

Issues at the district level are more difficult to identify with confidence. Some districts contain a low enough number of tweets that a few enthusiastic individuals can skew the results of the plot by tweeting prolifically using certain keywords. This produces unexpected results, such as the conservative Texas 31district tweeting “shutdownthegop” in much larger proportion than the liberal Texas 35district in Austin and San Antonio.

The inconsistent mix of tweet frequencies in districts causes some districts to almost always appear when the user clicks the button to show “different” word use. Florida’s 12th district is one of these. A close examination of the data shows the keyword “cruz” (in reference to Republican Senator Ted Cruz) appearing in 335 of the 471 tweets, likely the work of a small number of individuals. This effect could be mitigated by filtering out (or including only a sample from) users whose contributions exceed a particular percentage of the total tweets from the district.

The rankings of similar keyword patterns provide an interesting spatial indication of which units emphasized similar themes of conversation. Sometimes, these patterns match other spatial trends throughout the U.S. For example, clicking the “similar” word use button for the relatively religious western state of Utah yields a result that includes states in and along the “Bible Belt” in the southern U.S. All 10 of Utah’s most similar states have Republican Cook PVIs, a result that might be expected given Utah’s heavily Republican Cook PVI of R+22. In contrast, the similar states for Massachusetts (whose Cook PVI is Democratic leaning at D+11) are 80% Democratic, although they are spread out across the country (Figure 9). Some patterns are more difficult to explain, such as the moderately Republican leaning state of Georgia and Democratic leaning state of New York showing up in each other’s list of 10 most similar states.

In summary, explainable patterns are abundantly visible in SPoTvis, but they are not readily predictable for any pair of units. Adding more tweets for analysis would smooth out some of the most anomalous keyword usage patterns, especially in the congressional districts.

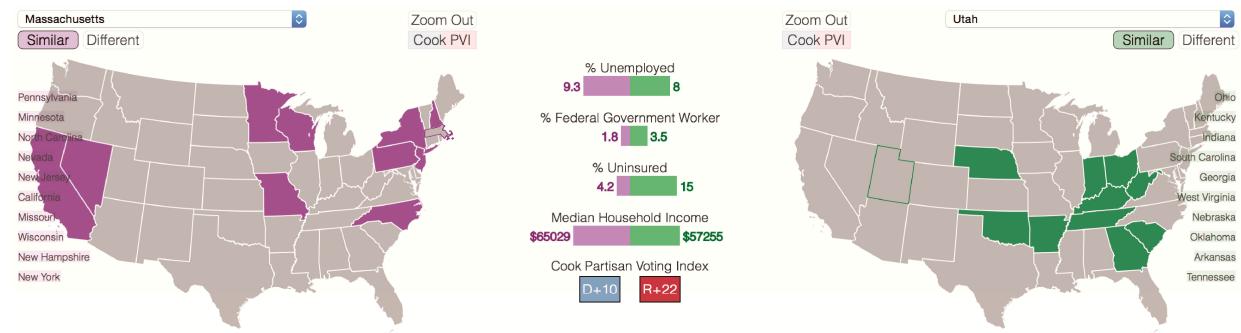


Figure 9. The set of states with similar keyword use sometimes follows political and cultural lines. Similar states to Massachusetts are shown on the **left**. Similar states to Utah are shown on the **right**.

5. SPoTvis User Evaluation

“The purpose of visualization is insight”.

-Stuart Card, Jock Mackinlay and Ben Shneiderman [27]

In this section, we explore the meaning of insight and report on various approaches to collecting insights in the context of a two-part user study of SPoTvis. In the first part of the study, participants were asked to choose roles for themselves (e.g., politician, political scientist, journalist, *etc.*). Based on those roles, participants were tasked with using SPoTvis to explore the Twitter data surrounding the government shutdown and to work through analytical tasks relevant to the roles chosen. Participants were asked to document insights they discovered and the approaches they took in obtaining those insights from using the tool. The second part of the study was an online survey designed to evaluate user experiences, interface design, functionality and future applications of SPoTvis to other contexts.

We first discuss the ways in which SPoTvis enabled users to answer and explore various questions and the processes they took to arrive at answers or to generate new hypotheses. We then evaluate the effectiveness of SPoTvis based on its ability to provide users with the necessary mechanisms for insight discovery. We conclude with promising directions for the future development and application of SPoTvis.

5.1. Contextualizing Insight

The goals of insight are discovery, decision-making and explanation. The usefulness of information visualization can be evaluated in terms of the extent to which such cognitive activities are achieved [27]. Formally defining insight, however, is problematic, because it takes many forms and is open to varying interpretations across and within disciplines that aim to measure it. Insight derives from the complexity of a dataset’s entirety. It accumulates and is generative. It is inexact, uncertain and qualitative in nature. It is unexpected, unpredictable and creative. The meaning of insight is relevant and embedded in the data domain [28]. Insight can be experienced in a spontaneous, indescribable and unrepeatable moment or characterized as a unit of discovered knowledge [29].

Because of the ambiguous meaning of insight, measuring a visualization’s ability to achieve its characteristics is challenging. Controlled experiments on benchmark tasks are a primary method for

evaluating visualizations, but force users to discover shallow, researcher-defined insights in short, definitive amounts of time [28]. Users are often asked to answer simple questions that can easily be used to measure the accuracy of the user's discovery of the predefined insight. This approach to measuring insight limits the unexpected, deep, qualitative and relevant insights the researcher might more readily obtain by allowing users to explore the visualization on their own and discover their own insights. The trade-off, of course, is that unconstrained exploration is likely to vary dramatically across individuals, and generalizations are difficult or impossible to derive from such unstructured activities. In the research presented here, we adopt a middle-ground using a semi-constrained exploration activity designed to provide opportunities for a wide range of insights on the part of participants, while providing us with a framework for synthesizing results.

5.2. Study Design

The purpose of visual analytics is to enable insight discovery through interactive visual interfaces [1]. The primary goal of conducting a user study on SPoTvis was to assess its ability to enable individuals with relevant expertise to discover their own spatio-political insights from the Twitter data surrounding the government shutdown. An additional aim of the study was to provide us with input on the design, functionality and future applications of SPoTvis. Thus, we designed our study in a way that captured the creative, deep and qualitative insights of users, as well as recorded straightforward answers to questions that assessed users' understandings of the tool's design and components. The study had two stages, outlined below.

The first stage asked participants to interact with SPoTvis and explore the Twitter data it accesses on their own within a week's time at their own pace. Participants were emailed a web link to SPoTvis, a link to an online demonstration video and a basic guide describing the tool and how to use it. Participants were first asked to create roles for themselves (e.g., politician, political scientist, journalist, etc.). Based on the selected roles, participants were tasked with writing short essays documenting interesting findings or patterns that they uncovered in the Twitter data, as well as the approaches they took to achieve those insights.

The second part of the assessment was an online survey. Participants were asked specific questions about their experiences using SPoTvis, SPoTvis' graphic design (the overall appearance of views and layout), interface functionality (in terms of both ease of use and ability to obtain information) and the applicability of the tool to other contexts.

Ten participants engaged in the study; nine completed both parts, while one participant chose to only take the survey. All participants had taken one or more courses in cartography, and the majority of participants actively do research in cartography, GIS and visualization. Three participants were graduate students recruited from the Department of Geography at the Pennsylvania State University. One participant was recruited from a Massive Open Online Course (MOOC) on mapmaking. Six participants were European academics recruited from the Department of Geoinformation Processing at the University of Twente in the Netherlands. The following sections present the results from the two-part user assessment of SPoTvis.

5.3. SPoTvis Assessment/Part 1

The first part of the SPoTvis user study resulted in nine reports, detailing tasks, approaches, types of interactions and obtained insights from varying roles or analytical perspectives, chosen by the participants. In the subsections below, we summarize report characteristics and the methods applied to analyze them, then present findings.

5.3.1. Report Analysis

Report length ranged from 147 to 511 words, with a mean word count of 375. The participant with the shortest report provided tables and screen shots to illustrate her findings and interactions with SPoTvis. Three participants assumed the roles of some type of journalist, two chose to be political scientists, one a state politician, one a human scientist and one a graphics enthusiast, and one did not specify a particular role (while these are adopted roles rather than characterizations of participants' actual expertise, for clarity of reporting below, we refer to the participants by their role). Tasks ranged from as vague as "exploration" to specific problems, such as, "How population composition (mainly Latin-American) affects the perception of the 2013 shutdown."

Reports were analyzed primarily by the lead author. *A priori* codes were designed to extract participants' tasks, types of interactions and achieved insights. Thus, we were particularly interested in understanding and synthesizing the workflows of SPoTvis' users, starting with their problem definitions and tracing the interactions that allowed them to arrive at solutions. Additionally, participants' reactions towards SPoTvis' design and functionality were flagged and related to the effectiveness of the tool in enabling insight discovery.

During the coding process, "analytical approach" emerged as a broad, organizational theme within which participants' workflows fit neatly. Four participants took data-driven approaches to analysis (*i.e.*, comparisons were made based on term usage, political or demographic questions); four participants took spatially-driven approaches (*i.e.*, comparisons were made based on places of interest or spatial questions); and one participant took a tool-driven evaluative approach to outline the (dis)advantages of SPoTvis for a specific context. In the following results subsection, findings from the reports are organized based on these three approaches.

5.3.2. Report Results

Participants who took data-driven approaches to analysis, in comparison to other participants, had more specific tasks they wanted to accomplish, engaged in more types of interaction and arrived at more specific insights (Table 3). Two of these participants chose to be journalists, one chose to be a human scientist and one a political scientist. These participants all had refined analytical objectives that extended beyond data exploration. The political scientist, for example, wanted to understand the relationship between ethnicity and perceptions of the government shutdown trending on Twitter. To investigate this relationship, the participant first referenced auxiliary ethnicity maps of the U.S. to better understand the spatial distribution of Hispanics and Latin Americans across the country. With this knowledge, the participant then selected New Mexico, the state with the highest Hispanic/Latino population composition, as the reference state and compared term usage between New Mexico and all

other states. The participant concluded that, “while a slight pattern in the similarity of terms can be seen in the states with highest ratio of Hispanic population, similar results appear in some of the states with the lowest Hispanic population. In general, the level of perception recoded by this dataset seems to reflect more the political preferences”. The political scientist’s workflow illustrates the insight-enabling process supported by SPoTvis and exemplifies the rich, qualitative and accumulative insights obtainable by users.

The data-driven group also leveraged SPoTvis’ more advanced functionality to support their sophisticated analytical reasoning. For example, the journalist and political scientist used the term demotion functionality to analyze relationships between political leanings and very specific subsets of terms. The Chinese journalist analyzed differences between term use and political party across scale using the “show similar” functionality and inferred that Democratic-leaning states (as an aggregate) tended to discuss (and potentially) blame the GOP, while Republican-leaning states (as an aggregate) focused more on Harry Reid in their discourse. At the individual district level, the Chinese journalist found that term usage varied regardless of party affiliation. Overall, SPoTvis’ functionality empowered the data-driven users to explore complex questions deeply.

The participants who took a spatially-driven approach to using SPoTvis assumed the roles of political scientist, university news journalist, graphics enthusiast and unspecified. As a group, these participants had both exploratory and more specific tasks they wanted to accomplish (Table 4). They framed their comparisons using known geographical boundaries, places of interest and spatial adjacency. The political scientist and news journalist sought to answer questions grounded in specific places of interest, comparing term usage, political leanings and demographics within defined spatial bounds. The graphics enthusiast and unspecified analysts pursued an unstructured exploration, documenting more-or-less random and disconnected findings.

In contrast with the data-driven analysts, these four participants often did not use SPoTvis’ more advanced functionality, such as term demotion or “show similar/different” spatial entities; nor did these users comment as frequently on potential links between term usage/political leanings and demographic variables, such as percent unemployed, percent federal worker, *etc.* However, although they used SPoTvis in a more restricted way, the spatially-driven users did generate some valuable insights. In comparison to the data-driven analysts, these analysts clearly had broader exploratory tasks and arrived at insights in a qualitatively different way. What these findings reveal is that SPoTvis effectively supported different approaches to analytical reasoning across varying roles and vastly different task definitions.

The last participant interpreted the first part of the SPoTvis user study in a slightly different way; in that the participant chose to evaluate SPoTvis based on the theoretical needs of a state politician. Rather than taking an analytical approach to explore the data and arrive at insights, the participant commented on the advantages and disadvantages of SPoTvis in meeting the needs of a state politician. The participant found that SPoTvis was able to quickly guide the politician in finding important topics within and across places and intuitively linked demographic variables to prominent topics. The politician, however, felt limited by a fixed (rather than user-specified) set of terms and overwhelmed by the use of color in showing multiple attributes. We consider these (and other) evaluative responses more thoroughly in the following section, which reports on the SPoTvis user experience, the effectiveness of the tool’s design and functionality and future applications of the tool to other contexts.

Table 3. Data-driven approaches to insight discovery.

Role	Task(s)	Approach	Interactions	Exemplary Insights
Chinese Journalist	Report to China on the 2013 U.S. Government shutdown	Data-driven	Compare all spatial entities by political party; compare two districts (same party); compare two districts (different parties); show similar districts	Democratic-leaning states blame GOP; Republican-leaning states likely blame Harry Reid; term usage varies regardless of party affiliation at the district level
Journalist	Find greatest difference in term use by political party; explore relationships between Cook PVI and demographics	Data-driven	Demote words to focus on specific subsets; compare all Democratic-leaning states/districts with all conservative-leaning states/districts using drop-down menus; click on map views to compare the spatiality of within/between political leanings with demographics	Terms “obama” and “gop” were more popular in Republican and Democratic-leaning states, respectively; terms “gopshutdown” or “shutdownthegop” were more often used in Democratic districts; terms “work”, “house”, “school” and “pay” correlated with districts having high levels of unemployment, uninsured and low median household income; in many states, the higher the Cook PVI was, the better welfare and income was
Human Scientist	Explore the concerns of population beyond the noise created by certain political terms	Data-driven	Demote “noise” words, such as “Obama”, “GOP”, “Democrats”, “Republicans”, etc.; compare remaining words and demographic variables between one Republican state (WY) and all Democratic-leaning states; demote “meaningful” words to analyze “noise” words by political party	Terms “pay”, “blame”, “school”, “talk” and “cnn” were used more by conservative leaning entities; terms “hope”, “debt”, “good”, “bad”, “furlough” and “worker” were used more by Democratic-leaning entities; terms “families”, “military” and “food” were used similarly between parties; Republican-leaning states used “noise” terms more frequently
Political Scientist	Explore the relation of ethnic origin (Hispanic) and the perception of the 2013 shutdown	Data-driven	Select NM as the reference state, because of its high Hispanic/Latino population composition; compare term usage between reference state and all other states	The pattern in the shutdown perception and ethnic origin is not clear; a slight pattern in the similarity of terms occurs in states with the highest ratio of Hispanic population, but similar results also appear in some of the states with the lowest Hispanic population; the level of perception recoded by this dataset seems to reflect more the political preferences of places

Table 4. Spatially-driven approaches to insight discovery.

Role	Task(s)	Approach	Interactions	Exemplary Insights
Political Scientist	Identify districts in Colorado (CO) that most closely reflect the Twitter behavior of the entire state	Spatially-driven	Compare the spread of keywords for each state-district pairing	CO 1 district is most similar in term usage to the aggregate view of the entire state of CO; CO 6 is most different in term usage to the aggregate view of CO
University News Journalist	Explore how congressional districts adjacent to the one that the University resides in compare in overall characteristics and term usage surrounding the government shutdown	Spatially-driven	Compare the spread of keywords for each university district-adjacent district pairing	Cook PVI is largely Republican for all districts in the study, yet large differences in term usage, potentially influenced by Democratic-leaning individuals associated with the university; terms populating nearby districts included “furlough”, “worker”, “employee”, “cost”, “boehner” and “money,” while university district used terms “cancel”, “food”, “washington”, “service”, “barackobama” and “gop” more often
Graphics Enthusiast	Exploration	Spatially-driven	Compare the spread of keywords between two states based on adjacency; compare the spread of keywords between districts within a state; compare the spread of keywords between the district and its respective state	OR in comparison with ID shows a clearer blame game as compared to FL in comparison with GA; polarized use of words between conservative MN 7 district and Democratic MN 8; Democratic TX 28 more focused on “congress” as compared to the overall conservative leaning of TX, which was more focused on “gop” and “obama”
Not Specified	Exploration	Spatially-driven	Compare the spread of keywords between two states based on adjacency; compare the spread of keywords between two states based on political leaning	WA and OR were seemingly interested in very different aspects of the shutdown; OR was more focused on the shutdown itself (GOP and members of the house), while WA expressed more opinion about potential reactions due to the shutdown (like work, money, debt, military, etc.); Twitter conversation gravitated around Obama and Obamacare in conservative-leaning states, while Democratic leaning states conversed more about the GOP and Republicans

5.4. SPoTvis Assessment/Part2

The second part of the SPoTvis user study was an online survey consisting of a Likert scale, multiple choice and short answer questions designed to collect feedback on the SPoTvis user experience, the overall appearance of SPoTvis' views and layout, interface functionality and applicability of the tool to other contexts (Appendix I). Ten participants completed the survey, nine of which also completed the first part of the study. Six participants reported having spent over an hour using SPoTvis; one spent 30–45 min; and three spent 15–30 min. All participants recorded having watched the demonstration video prior to completing the survey.

5.4.1. User Experience

The first part of the survey aimed to assess users' reactions to and feelings toward SPoTvis (Figure 10). Three users found their initial experiences using SPoTvis easy and fluid. Four participants found their initial experiences neither confusing nor particularly intuitive, while three users initially felt the tool was more confusing than easy to use. Nine participants felt positively towards SPoTvis while they interacted with it, while one participant felt neutrally towards the tool. Moreover, most users commented on the interface being easy to use once accustomed to its functionality.

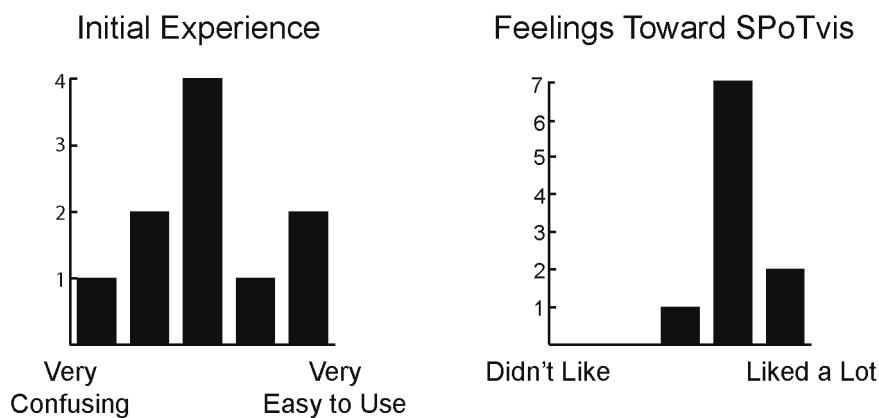


Figure 10. Initial experiences (**left**) and feelings toward SPoTvis (**right**). Each graph depicts the frequency of response on a five-step Likert scale with the end points labeled using the terms shown below the x-axis.

5.4.2. Design Evaluation

Questions in the SPoTvis design portion of the survey aimed to evaluate the appearance of SPoTvis' views and layout. The first questions asked users to rate the effectiveness of the tool's design, how aesthetically pleasing they found the design to be and how effective the use of color was in the design (Figure 11). While most participants reacted positively toward the overall design and aesthetics, responses varied on the effectiveness of color in the design. More specific reactions to color use in SPoTvis are discussed below.

Following the design rating questions were short answer questions that first assessed users' understandings of design choices, then allowed users to comment on aspects of the tool they found to be

most innovative and most confusing. Participants were also encouraged to provide design suggestions in this part of the survey.



Figure 11. Ratings on SPoTvis' design (**left**), aesthetics (**middle**) and use of color (**right**). Each graph depicts the frequency of response on a five-step Likert scale with the end points labeled using the terms shown below the x-axis.

When asked what the colors of bubbles represented in the term polarity plot, six respondents correctly linked the shades of purple and green to their respective map views, depicted below the plot. However, two respondents confused bubble color with political leanings, rather than spatial comparisons. The remaining two participants commented on the use of color more broadly instead of directly answering the question of what color represents in the plot. Many users felt that SPoTvis overused color to represent too many attributes. SPoTvis uses color to encode left/right comparisons, political leanings, similarities, differences and combinations of these via differing fill, stroke and on-hover highlight colors. This explains some of the confusion participants had in understanding what the color of bubbles represented in the term polarity plot, particularly in the context of using SPoTvis for relatively short periods of time.

When asked what bubble/font size represented in the term polarity plot, seven respondents completely understood that the visual variable of size encoded the frequency of term usage between two spatial comparisons. The other three respondents did not necessarily misunderstand what bubble/font size conveyed, but rather provided responses on how sizes should have varied more (or less) or simply stated that “size worked”.

Aspects of the design participants found most innovative included: the term polarity plot, term polarity plot coordination with map views, integration of bubble placement with spatial comparisons, left/right aspect for comparison, multiscale visualization and bubble movement. One participant commented, “I think including the plot reacting to the interaction with the maps is something innovative in the geovisualization field”. The participant found the separated approach to information design effective in reducing cognitive overload, thus making the data more interpretable. Overall, reactions focused on the clean design of the term polarity plot and comparison views, the core and most unique components of SPoTvis. More subtle suggestions participants provided for refining the design of SPoTvis included: labeling the term polarity plot’s x-axis and adding map labels to provide a better sense of place.

5.4.3. Functionality Evaluation

The SPoTvis functionality section of the survey evaluated how easy SPoTvis was to use, as well as its ability to provide users with insights. Eight participants found SPoTvis' interaction options obvious and useful, while two participants felt neutrally towards the tool's overall functionality (Figure 12). Because comparison was central to the design of SPoTvis, we were particularly interested in using this section of the survey to understand how users made comparisons. In response to a question about how spatial comparisons were made (when presented with three, multiple choice options), six participants reported that they used both drop down menus and map clicking (on states/districts) to make comparisons; two participants used only drop down menus; and two participants used only map clicking.

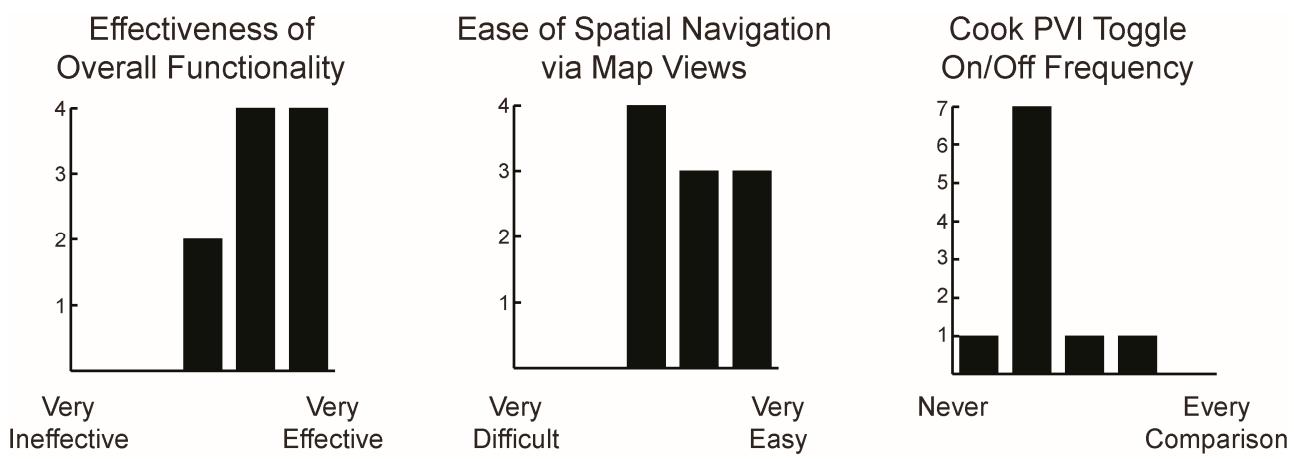


Figure 12. Ratings on SPoTvis' functionality (**left**), ease of spatial navigation (**middle**), and Cook PVI toggle frequency (**right**). Each graph depicts the frequency of response on a five-step Likert scale with the end points labeled using the terms shown below the x-axis.

With respect to map use and interaction, users had mixed reactions to spatial navigation using the map views. Six participants were able to easily navigate in the map views, while four participants assigned a neutral/intermediate rating to SPoTvis navigation. One participant specifically commented on how map navigation supported insight gathering: “It is good you classified the zoom levels in three stages. The user is never out of the scope of the map, and the feature under study is always highlighted. And it is easy to go up and down between levels”. Two participants suggested adding “pan/zoom” functionality to enhance way finding within the map views. Eight of the ten participants found the Cook PVI base map layer very useful for linking political leanings to spatial comparisons. These participants chose to almost never toggle the layer off. The other two participants found it useful to occasionally switch the layer on and off during exploration.

To evaluate functionality in the term polarity plot, we asked users to describe in short sentences what bubble and term movement/placement represented. Eight participants correctly inferred that movement and placement of bubbles/terms were a function of the combined, relative use of the terms between two spatial comparisons. The other two participants misinterpreted movement and placement in the plot as being related to political party or tweet composition, rather than to spatial units. When asked how SPoTvis' coordinated functionality supported insight gathering, participants commented on how intuitive it was to make comparisons using the term polarity plot together with the map views.

One participant, for example, stated: “The speed of clicking in the maps and getting the new plot ready in (a) few seconds is quite good, because you never get tired of exploring the plot.” For this user, interaction inspired additional exploration. Another user commented, “There was enough flexibility in the interactions to support experimentation, which resulted in new insights”. Participants further found the term demotion functionality useful for analyzing very specific term subsets of interest. Overall, results indicate that SPoTvis’ interaction options engaged users and fostered data exploration.

When asked how interaction detoured/inhibited insight gathering, participants felt limited by the ability to only compare two entities or two groups of entities at a time. Participants further wanted a “history of interactions” to more rigorously collect and maintain acquired insights. The “similar/different” functionality was critiqued on the basis that users wanted a better explanation of what the similarity measure took into account, as well as a more complex similarity criteria. For example, one participant commented: “I was expecting, say, to find states or districts similar to others in terms of the four parameters (demographic variables) between both maps”. Currently, SPoTvis derives similarity only on the basis of how spatial units used key terms. Extending the SPoTvis capability to explore similarity in multiple ways and using multiple criteria together is a useful suggestion that we plan to explore in future versions of SPoTvis.

5.4.4. Future Applications and Summary

In the final part of the survey, we asked participants to envision ways in which SPoTvis might be used in someone else’s work. Participants provided very interesting and highly diverse suggestions that ranged from very broad in scope (e.g., “any exploratory analysis”) to specific applications in crisis management, economics, environmental studies, journalism, law, political science, social science and technical science. One participant imagined SPoTvis “being useful to do comparative sightings between places in wildlife scenarios, protest landscapes and mapping movement patterns.” Another participant envisioned SPoTvis “as a generic near real-time tool for human pattern analysis”. This participant went on to say that SPoTvis “could help to find patterns tied to (the) geographic domain, probably never represented before as a combination of volunteered geographic information (VGI) and maps, that can be used as guidelines or triggers for other ideas”. One participant responded to the question from a geovisual analytics tool development perspective, commenting on the benefit of integrating the D3 visualization framework into more traditional geographic information systems to advance spatial data exploration and reasoning. Lastly, one participant felt that visualization teachers might want to demonstrate SPoTvis to their students to illustrate innovative design in the geovisualization field. Clearly, there exist many exciting opportunities to extend SPoTvis to other research areas, as well as to integrate the tool’s novel functionality into more comprehensive spatial analysis software.

The SPoTvis user experience was generally positive and enjoyable. Users appreciated the aspirations, fluidity and novelty of SPoTvis’ design and functionality. In some instances, SPoTvis’ options for interaction clearly inspired data exploration. The term polarity plot and left/right comparison views were considered by participants to be valuable contributions to support the analysis of perspectives. Users further provided useful suggestions for improving the usability of SPoTvis in future versions, with an emphasis on how color is used and on improvements for the “similar/different” functionality. Moving forward, we plan to explore the use of texture, as suggested by one participant, to alleviate the cognitive

overload associated with encoding too many attributes by color. For example, we could represent similar and different spatial units on the map using varying densities of areal patterns. We also plan to develop a more sophisticated measure of similarity, one that is more intuitive and takes into account not just similarities in shared term usage, but also in political leanings and demographic variables. The careful addition of axis and place labels will also be considered to more directly inform users about the variables' meanings and to provide a better sense of place in the map views.

A key strength of the design and functionality of SPoTvis is that it is tailored to a special purpose, bounded in time, space, observations and context. Our longer-term ambitions for the future development of geovisual analytics applications will focus on extending SPoTvis' functionality to be more flexible and scalable. We envision a SPoTvis-inspired application that consumes user-specified data of any applicable types in real-time from streaming feeds (not just Twitter, but photo tags, RSS news feeds and others), allowing analysts in any number of disciplines to spatially compare trending or most relevant terms, topics or images in the term polarity plot. The core views and interactions unique to SPoTvis would remain, but analysts would have control over selecting an appropriate number of entities to assess in the term polarity plot and the ability to aggregate their data across meaningful areal units.

6. Conclusions

Although the above trends and results are interesting, do they tell us anything new? Spatial patterns of political opinions and voting behavior in the U.S. are well documented, but always changing. Our analysis and tool demonstrate that people's conversations on social media often reflect these known trends and are worth monitoring to detect new or shifting patterns. It seems feasible to suggest that applications like SPoTvis may be employed more often in the future to act as real-time forecasters of political winds.

We identified 50 of the most common words and topics in tweets generated in fall, 2013, related to the U.S. Government shutdown and ACA implementation. We observed that these keywords touched political figures, the labor force and domestic life. Furthermore, we identified that areas with similar political persuasion tended to tweet about a similar set of topics.

At the same time, it is important to note what people are not talking about. For example, various news stories reported the closure of national parks and museums due to the shutdown, but these topics did not appear in the list of most common words in the tweets we examined. News stories about children being locked out of the zoo may provide dramatic imagery for the media, but we do not find evidence that people are consistently discussing these topics nationwide to the degree that they are discussing certain other subjects (like impacts on their jobs or the politics of the shutdown).

The results shown in SPoTvis represent aggregations that may not appear when using other scales or areal units. At the individual level, many people's political opinions and tweets do not strictly follow party lines, just as at the state level, not every voter chooses the same party. There are plenty of Democrats in Texas or Republicans in Massachusetts whose local influences are underrepresented when examined at the state or national scale. This was a reason we wanted to include congressional districts on the map despite the relatively low number of tweets. Furthermore, party preferences along urban-rural or ethnic divides are not well represented by state lines, while congressional districts may intentionally be drawn to follow these boundaries.

Our research might be strengthened by an increased number of tweets and a wider variety of tweet contributors. It is important in any analysis that uses social media as primary data to consider the biases in these data related to both population bias (e.g., who has access to the Internet and who is most inclined to use these media) and sampling bias of algorithms and application programming interfaces through which researchers are allowed to sample the data [30]. We suspect that the strength of the relationship between a unit's Cook PVI and that of its most similar and dissimilar units would be increased if there were more tweets in the dataset, but that there would still be bias due to the difference between the population that the PVI measures and that reflected on Twitter. The addition of some "smart filtering" to detect the influence of overly dominant tweeters in the dataset could also help make the results more meaningful.

An increased number of tweets in the dataset would allow for temporal analysis, examining how themes in conversation changed as the shutdown was approached, endured and resolved. The fluidity of movement offered by the bubbles in the term polarity plot could create an ideal environment for showing how topics shifted in prominence and place of origin during consecutive time periods. The temporal information is readily available in the timestamps of the existing tweet data; however, there are probably not enough tweets in the dataset examined here to create any meaningful temporal representations at the congressional district level; however, state-level aggregates per day might have high enough frequencies to result in usable data.

Acknowledging the above limitations, we assert that SPoTvis successfully addresses many of the challenges set forward by Elwood [31] for dealing with heterogeneous, qualitative and dynamic spatial data. For example, SPoTvis supported deep, qualitative insight gathering by nine analysts who adopted highly diverse roles seeking to address distinctly different tasks. SPoTvis' design and functionality allowed users to take both data-driven and spatially-driven approaches to exploring, generating new hypotheses from and making inferences about the discourse posted on Twitter surrounding the 2013 U.S. Government shutdown. The interaction between the map and the term polarity plot to compare any two spatial units is a unique visual analytics technique that could be adapted to a variety of datasets and scales, such as RSS news feeds from across the world. Ultimately, SPoTvis succeeds at mining a set of personal expressions via tweets and demonstrating how the themes therein are connected to known patterns of political persuasion in the United States.

Supplementary Materials

Supplementary materials can be accessed at: <http://www.mdpi.com/2220-9964/4/1/337/s1>.

Acknowledgments

The authors wish to thank Marcel Salathé, Zhuojie Huang and Todd Bodnar for assistance with collecting the tweets, Maggie Houchen for her feedback on early iterations of this research and the reviewers for their time and invaluable feedback. This work was supported by the National Science Foundation under IGERT Award #DGE-1144860, Big Data Social Science and Pennsylvania State University.

Author Contributions

Jonathan was the lead designer for SPoTvis, processed spatial data, assisted in coding the application, conducted and analyzed data from the user study and participated in the project from start to finish. Sterling provided project management and design guidance for SPoTvis, as well as contributed to the programming of the back-end data infrastructure. Brian was the lead programmer for SPoTvis. Wanghuan processed raw tweets and conducted exploratory and statistical analysis of the tweets. Alan advised every aspect of the project, contributed to design of the user study and helped recruit participants for the study. All authors contributed to writing the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Thomas, J.J., Cook, K.A., Eds. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, 1st ed.; IEEE Computer Society: Los Alamos, CA, USA, 2005.
2. MacEachren, A.M. Cartography as an academic field: A lost opportunity or a new beginning? *Cartogr. J.* **2013**, *50*, 166–170.
3. GeoVISTA “SPoTvis”. Available online: <http://www.geovista.psu.edu/SPoTvis/> (accessed on 14 February 2015).
4. Vieweg, S.; Hughes, A.L.; Starbird, K.; Palen, L. Microblogging during two natural hazards events: What twitter may contribute to situational awareness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; pp. 1079–1088.
5. Signorini, A.; Segre, A.M.; Polgreen, P.M. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS One* **2011**, *6*, e19467.
6. Scharl, A.; Hubmann-Haidvogel, A.; Weichselbraun, A.; Lang, H.P.; Sabou, M. Media watch on climate change—Visual analytics for aggregating and managing environmental knowledge from online sources. In Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS), Grand Wailea, HI, USA, 7–10 January 2013; pp. 955–964.
7. MacEachren, A.M.; Robinson, A.C.; Jaiswal, A.; Pezanowski, S.; Savelyev, A.; Blanford, J.; Mitra, P. Geo-twitter analytics: Applications in crisis management. In Proceedings of the 25th International Cartographic Conference, Paris, France, 3–8 July 2011; pp. 3–8.
8. Smith, M.A.; Rainie, L.; Shneiderman, B.; Himelboim, I. Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters. *Pew Intern. Am. Life Proj.* 2014. Available online: <http://www.pewinternet.org/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/> (accessed on 16 September 2014).
9. Diakopoulos, N.; Naaman, M.; Kivran-Swaine, F. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (VAST), Lake City, UT, USA, 24–29 October 2010; pp. 115–122.
10. Xu, P.; Wu, Y.; Wei, E.; Peng, T.Q.; Liu, S.; Zhu, J.J.; Qu, H. Visual analysis of topic competition on social media. *IEEE Trans. Vis. Comput. Graphics* **2013**, *19*, 2012–2021.

11. Havre, S.; Hetzler, B.; Nowell, L. ThemeRiver: Visualizing theme changes over time. In Proceedings of the IEEE Symposium on Information Visualization, Salt Lake City, UT, USA, 9–10 October 2000; pp. 115–123.
12. Tsou, M.H.; Yang, J.A.; Lusher, D.; Han, S.; Spitzberg, B.; Gawron, J.M.; Gupta, D.; An, L. Mapping social activities and concepts with social media (Twitter) and web search engines (Yahoo and Bing): A case study in 2012 US Presidential Election. *Cartogr. Geogr. Inform. Sci.* **2013**, *40*, 337–348.
13. Bodnar, T.; Salathé, M. Validating models for disease detection using Twitter. In Proceedings of the 22nd International Conference on World Wide Web Companion, Rio de Janeiro, Brazil, 13–17 May 2013; pp. 699–702.
14. Viégas, F.B.; Wattenberg, M. Timelines tag clouds and the case for vernacular visualization. *Interactions* **2008**, *15*, 49–52.
15. Funkhouser, H.G.; Walker, H.M. Playfair and his charts. *Econ. Hist.* **1935**, *3*, 103–109.
16. IBM “Many Eyes”. Available online: <http://www-969.ibm.com/> (accessed on 5 April 2014).
17. Viegas, F.B.; Wattenberg, M.; van Ham, F.; Kriss, J.; McKeon, M. Manyeyes: A site for visualization at internet scale. *IEEE Trans. Vis. Comput. Graph.* **2007**, *13*, 1121–1128.
18. Wordle. Available online: <http://www.wordle.net/> (accessed on 5 April 2014).
19. Viegas, F.B.; Wattenberg, M.; Feinberg, J. Participatory visualization with Wordle. *IEEE Trans. Vis. Comput. Graph.* **2009**, *15*, 1137–1144.
20. Cui, W.; Wu, Y.; Liu, S.; Wei, F.; Zhou, M.X.; Qu, H. Context preserving dynamic word cloud visualization. In Proceedings of the Pacific Visualization Symposium (PacificVis), Taipei, Taiwan, 2–5 March 2010; pp. 121–128.
21. Cui, W.; Liu, S.; Tan, L.; Shi, C.; Song, Y.; Gao, Z.; Qu, H.; Tong, X. Textflow: Towards better understanding of evolving topics in text. *IEEE Trans. Vis. Comput. Graph.* **2011**, *17*, 2412–2421.
22. Collins, C.; Viegas, F.B.; Wattenberg, M. Parallel tag clouds to explore and analyze faceted text corpora. In Proceedings of the 2009 IEEE Symposium on Visual Analytics Science and Technology (VAST), Atlantic City, NJ, USA, 11–16 October 2009; pp. 115–122.
23. Bostock, M.; Carter, S.; Ericson, M. At the National Conventions, the Words They Used. *New York Times*, 2012. Available online: <http://www.nytimes.com/interactive/2012/09/06/us/politics/convention-word-counts.html/> (accessed on 13 December 2013).
24. Harrower, M.; Brewer, C.A. ColorBrewer.org: An online tool for selecting colour schemes for maps. *Cartogr. J.* **2003**, *40*, 27–37.
25. Shneiderman, B. The eyes have it: A task by data type taxonomy for information visualizations. In Proceedings of the 1996 IEEE Symposium on Visual Languages, Boulder, CO, USA, 3–6 September 1996; pp. 336–343.
26. Bostock, M.; Ogievetsky, V.; Heer, J. D³ data-driven documents. *IEEE Trans. Vis. Comput. Graph.* **2011**, *17*, 2301–2309.
27. Card, S.K., Mackinlay, J.D., Shneiderman, B., Eds. *Readings in Information Visualization: Using Vision to Think*, 1st ed.; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1999.
28. North, C. Toward measuring visualization insight. *IEEE Comput. Graph. Appl.* **2006**, *26*, 6–9.
29. Chang, R.; Ziemkiewicz, C.; Green, T.M.; Ribarsky, W. Defining insight for visual analytics. *IEEE Comput. Graph. Appl.* **2009**, *29*, 6–9.

30. Ruths, D.; Pfeffer, J. Social media for large studies of behavior. *Science* **2014**, *346*, 1063–1064.
31. Elwood, S. Geographic Information Science: New geovisualization technologies—emerging questions and linkages with GIScience research. *Progress Hum. Geogr.* **2009**, *33*, 256–263.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).