

Visualizing Government Data on the Web

Peter Kinnaird

College of Computing/GVU Center
Georgia Institute of Technology
Atlanta, GA USA 30332-0280

ABSTRACT

The most critical watchdog on democracy in the 21st century is the dispersal of information. Citizens have unprecedented access to vast amounts of data regarding numerous governmental processes. These data are often prohibitively large and complex for gaining meaningful insights and understanding. Through this position paper, I advocate for the creation of web-based visualizations to handle these increasingly available government data. I describe some existing problems with understanding these data and then cite examples to demonstrate how Information Visualization can help. Finally, I discuss some of advantages of putting visualizations on the web as well as a working prototype of a web-based visualization using the latest browser technologies.

KEYWORDS: Information Visualization, Grand Challenges, Congress, Digital Governance, Internet Information Applications

INDEX TERMS: H.5.0 [Information Interfaces and Presentation]: General; J.0 [Computer Applications]: General

1 OBSTACLES TO UNDERSTANDING GOVERNMENT DATA

The *IEEE Infovis 2008 Grand Challenges Panel* [1] described a goal of “Total Political Transparency.” By providing civic oversight, we can increase civic participation and reduce government corruption. Even before the launch of Data.gov, an enormous amount of government data was available online. These data include detailed census data, Congressional voting records, the text of speeches given in Congress, and topographical maps showing flood probabilities. Although it is admirable that the government provide these data for citizen use and analysis, there have been three fundamental obstacles to civic use: (1) the data are exceedingly complex; (2) the data are often presented in an unnecessarily confusing format; and (3) the data have poor accessibility.

In an informal trial, it took two Computer Science graduate students over half an hour to find the detailed results of the Congressional vote for Obama’s proposed Stimulus Bill of January 2009. This example demonstrates all three of the aforementioned obstacles. First, the data is exceedingly complex. There were a large number of votes associated with the Stimulus Bill. These votes were primarily on amendments to the Bill and procedural votes to alter the status of the Bill. Some or all of these votes could be of interest to citizens. The confusing aspects of these data primarily concern terminology and process. For example, interested parties might not know if they are seeking a bill, a joint resolution, a concurrent resolution, or a simple resolution, let alone the differences between those.

* e-mail: peterkinnaird@gatech.edu

2 VISUALIZATION

Dealing with all of this data leads to a variety of possible techniques for gaining insights. Search certainly has its place and can be a very useful tool. However, I would argue that the domain of government data is ripe for visualization. For example, Death and Taxes[2] provides a unique look at the United States federal budget. A quick examination of this visualization enables users to gain insights that would otherwise be abstruse. The 2009 federal budget weighs in at over 5000 pages, certainly meeting all three of the stated obstacles: complex, confusing, and inaccessible.

Congressional voting records have already been demonstrated to meet all three of the obstacles. The SocialAction[3] system demonstrated that interactive graph visualization can be used to gain interesting and unintuitive insights into these data.

3 ON THE WEB

Putting government data into the hands of the people means not only removing all three of the obstacles discussed, but also diffusing the information as widely as possible. The web is the most accessible platform available. Additionally, the potential for including social networking capabilities alongside the visualizations opens up whole avenues of techniques for interacting with the data.

The latest generation of browsers is beginning to support HTML5. With HTML5 comes the *canvas* tag, enabling immediate mode 2D drawing similar to that found in Java and other programming languages. Improvements to browser speed and Javascript support enable the use of this tag to develop entire web applications. I have already succeeded in developing a prototype visualization system for Congressional voting records that uses these new HTML features. Although support is limited, it is improving with each new browser release, and performance is already surprisingly good. Renewed energy in the browser wars from the competitive releases of Chrome and Safari ensures that full support of the *canvas* is near.

4 CONCLUSIONS

Government data is complex, confusing, and inaccessible. It is a domain ripe for visualizations in order to enable civic oversight, particularly with the go-live of Data.gov. The Web is the obvious platform for diffusing these data. Improvements to browser support for HTML5 enable the development of visualizations for the web more easily and effectively than ever before.

REFERENCES

- [1] Bachman, J. <http://www.wallstats.com/deathandtaxes/>. Retrieved July 16, 2009.
- [2] Daniel Keim & Tamara Munzner. Grand Challenges for Information Visualization, *IEEE InfoVis*, May 2008.
- [3] A. Perer and B. Shneiderman, “Integrating Statistics and Visualization for Exploratory Power: From Long-Term Case Studies to Design Guidelines,” *IEEE Comput. Graph. Appl.*, vol. 29, 2009, pp. 39-51.