

No. 27
September 2008

MERCATUS ON POLICY

OPENING GOVERNMENT: MASHUPS

By Jerry Brito

MERCATUS CENTER
GEORGE MASON UNIVERSITY

MAKING GOVERNMENT INFORMATION available online would not only benefit individual users of government websites, it would also make it simpler for third parties to aggregate government data. By aggregating data, websites can present government information in innovative and useful ways. For example, federal spending data gathered from a government website could be presented by a third party as an interactive map that shows the locations of funding recipients. Such applications make data exponentially more valuable. Government need not develop such innovative tools itself; as long as the data is made available online in a structured format, private parties will make good use of it.

“Structured data” is a term of art. It means that information is presented in a format that allows computers to easily parse and manipulate it. While a static web page that lists a series of news stories or proposed regulations is not structured, the web page may have a companion XML file containing the same information. Structured XML allows a user to sort the data by ascending or descending date, alphabetically by headline or author, by number of words, and in many other ways that a static web page does not afford.

When the government makes data available in a structured format, it opens the doors to innovative and enlightening remixes of information known as mashups. Mashups are tools that can be used by journalists, bloggers, and citizens to better analyze and monitor government’s activities.

MASHUPS

THE TERM “MASHUP” has its origins in music. The advent of digital editing technologies made it relatively simple for DJs and amateurs to take two or more songs and mash them together to produce novel creations. The paradigmatic example of a music mashup may be Danger Mouse’s highly acclaimed and highly illegal “The Grey Album,” which mixed music from The Beatle’s “The White Album” with vocals from rapper Jay-Z’s “The Black Album.”

The term mashup now extends to applications that mix together disparate sets of data to create new and unique information. For example, the popular free classified ad site Craigslist.com is an almost definitive source for rental housing listings in urban areas. However, the site lists ads in the

Mashups built on open interfaces and structured data represent a great potential fount of information about the workings of government.

order users add them to the site. This means, using the Washington D.C. metro area as an example, one listing could be for an apartment in the Adams Morgan neighborhood of the District and the very next ad would be for a house in Alexandria, Virginia. This frustrated software engineer Paul Rademacher when he was looking for a place to live in Silicon Valley in 2004. So, he built HousingMaps.com, a mashup of the listings from Craigslist.com and Google Maps. This mashup allows users to bring up a map of the area in which they are interested (say five square blocks in a particular neighborhood). Then, pushpin icons will appear representing the properties available for rent in that area. Clicking on a HousingMaps.com pushpin brings up a bubble with the rental listing data including rooms, price, location, photos, and a link to the actual listing.

What is amazing about HousingMaps.com is it offers a new and unique information source that is richer, and more useful, than either Craigslist or Google Maps alone. What makes this possible is Google’s choice to make its maps application interface open for anyone to use and Craigslist’s similar choice to make its data available in an open and structured format.

These decisions to support openness and useful data formats allowed for an innovation that neither company could have predicted would emerge.

Mashups built on open interfaces and structured data represent a great potential fount of information about the workings of government. Mashups produce varied and unexpected outcomes that could make government activities more transparent, and reveal patterns currently hidden in murky mountains of unstructured data. To get a sense of what is possible, we can take a look at a leading transparency mashup called MAPLight.org.

The MAP in MAPLight.org stands for “money and politics,” and the site’s mission is to illuminate the connection between the two. Founded by computer expert Dan Newman, the site mashes together congressional voting data with campaign finance information. The result is a searchable database that highlights the connections between campaign contributions and how members of Congress vote. Using the MAPLight database, users can look up a particular bill and see the interest groups, as well as the individuals and corporations, who support and oppose it. MAPLight also allows users to look up individual members of Congress in order to see how they voted on a particular bill and to see how much money they received from groups supporting and opposing the bill. This is a new window into congressional actions that legislators did not previously need to consider. Such a mashup would not be possible without the structured data government often fails to provide and is being made accessible by third-party “hacks” from non-profit groups such as GovTrack.us and OpenSecrets.org.

Another mashup aimed at increasing government transparency is OpenCongress.org. Among other things, this site takes bill and vote data from GovTrack.us and mashes it with data feeds from blogs and mainstream news sources so that one can pull up a page for a bill or a legislator and see news stories and blog posts that mention the bill and/or legislator.

RECOMMENDATIONS

THE FOUNDATION ON which Internet technologies can help improve transparency is the idea, to the greatest extent feasible, government data should be made public. The next building block is the idea information should not just be made available online, but online resources must also be useful. This means putting data online in structured, open, and searchable formats.

Structured means the data is presented in a machine-readable format that makes it easy for individuals to subscribe to discrete data feeds and for others to use the data in their own creations—that is, as the source data for a community site such as WashingtonWatch.com or mashups like MAPLight.org.

[Getting Started](#)
[Latest Headlines](#)
[Apple](#)
[Amazon](#)
[eBay](#)
[Yahoo!](#)
[News](#)

[For Rent](#)
[For Sale](#)
[Rooms](#)
[Sublets](#)

City:
 Price:
[Show Filters](#)
[Refresh](#)

Powered by [craigslist](#) and [Google Maps](#)

[About/Feedback](#)

Drag the map with your mouse, or double-click to center.
 Click on an  icon to select a city.
 In each city, click on the  and  icons to see listings.
 You can also click the  and  icons next to each listing.
 Yellow icons have pictures.
 Set the price range using the drop-down menu above.



500 mi
 500 km

[Getting Started](#)
[Latest Headlines](#)
[Apple](#)
[Amazon](#)
[eBay](#)
[Yahoo!](#)
[News](#)

[MAPLight.org](#)

MONEY AND POLITICS: ILLUMINATING THE CONNECTION

[U.S. Congress](#)
[Interest Groups](#)
[Legislators](#)
[Bills](#)

[About Us](#)
[Press Room](#)
[Widgets](#)
[Participate](#)

[California Legislature](#)

> Current Bills in Congress

Paycheck Fairness Act
 Lily Ledbetter, Pay discrimination, Civil Rights Act

Family Smoking Prevention and Tobacco Control Act
 Philip Morris, Menthol Cigarettes

A bill to provide needed housing reform and for other purposes.
 housing, Fannie Mae, Freddie Mac, Foreclosure, Bankruptcy

Global Online Freedom Act of 2007
 China, Firewall

[SEE MORE](#)

> News About MAPLight.org

MAPLight.org is hiring - Drupal Developer/MySQL Admins needed

MAPLight.org honored to be one of eight recipients of the 2008 Knight-Batten Awards for Innovations in Journalism

San Francisco Chronicle: Site Tracks Influence of Interest Groups by Tom Abate

Media General News Service: Phone Giants Gave Liberally to Lawmakers During Wiretapping Debate by Neil Simon

Time Magazine: The Citizen Watchdogs of Web 2.0 by Jeremy Caplan

[SEE MORE](#)




Interest Groups


Legislators


Bills

MAPLight.org Widgets
 Get the latest Presidential Fundraising Widget.

Funds Raised	MAPLight.org
Barack Obama	\$431,311,980
John McCain	\$182,601,429

2008 Presidential Campaigns, last updated through Aug. 11, 2008. Source: FEC, September 8, 2008.

Comment of the Week
 Hysterics Over Hate Crime Legislation
 It's easy to get caught up in the impassioned rhetoric more...
 Posted by *Maya Perelman* on Aug 25, 2008

Hot Topics

- HR 3221 (110th) (4 comments)
- HR 5632 (110th) (3 comments)
- HR 5982 (110th) (3 comments)
- HR 5546 (110th) (2 comments)
- HR 168 (110th) (2 comments)
- HR 1092 (110th) (2 comments)
- HR 1834 (110th) (2 comments)
- S 3207 (110th) (2 comments)
- HR 2401 (110th) (2 comments)
- HR 1495 (110th) (2 comments)

[SEE MORE](#)

Recent Comments

Aviation Bill Exposes Need for Reform in FAA
 Citing the FAA's too-close relationship with the airlines it regulates more...
 Posted by *Maya Perelman* on Sep 4, 2008

Putting Environmentalism on the Farm Bill Agenda
 The lack of public support or opposition to the bill more...
 Posted by *Maya Perelman* on Sep 4, 2008

Bond Ratings
 This bill would require the SEC to make each credit more...
 Posted by *Ryan Birss* on Sep 4, 2008

Honoring Iraq War Veteran with Post Office Name
 Peter J. Navarro gave his life fighting in Iraq. To more...
 Posted by *Ryan Birss* on Sep 4, 2008

Honoring Veteran with a Post Office

Video Tour
 All about us in just 6 minutes!



Example Pages

Interest groups

- Accident & health insurance
- Teachers unions
- Casinos, racetracks & gambling

Legislators

- Hancy Pelosi
- John Boehner
- James Whalle

Bills

- Horse protection
- Omni free trade
- Labor membership and organizing
- Executive compensation
- Gambling
- Collective bargaining
- Prescription drugs
- Stem cell research

Open means that the digital formats chosen should be non-proprietary and widely accepted. Open formats are often created and maintained by independent standards organizations and are free of copyright restrictions on their use. For example, MP3 is an open audio file format, while RealMedia and Apple QuickTime are proprietary.

CONCLUSION

TO HOLD GOVERNMENT accountable for its actions, citizens must know what those actions are. To that end, they must insist government act openly and transparently to the greatest extent possible. In the twenty-first century, this entails making its data available online and easy to access. If government data is made available online in useful and flexible formats, citizens will be able to utilize modern Internet tools to shed light on government activities. Such tools include mash-ups, which highlight hidden connections between different data sets.

Today, however, the state of government's online offerings is very sad. Some nominally publicly available information is not online at all, and the data that is online is often not in useful formats. Government should release public information online in a structured, open, and searchable manner. To the extent that government does not modernize, however, we should hope that private third parties build unofficial databases and make these available in a useful form to the public.

ENDNOTES

1. Mark Doms and Meryl Motika, "The Rise in Homeownership," *Federal Reserve Bank of San Francisco Economic Letter* 2006-30 (November 3, 2006), <http://www.frbsf.org/publications/economics/letter/2006/el2006-30.html>.
2. Housing and Household Economic Statistics Division, U.S. Census Bureau, Housing Vacancy Survey, Third Quarter 2007, table 5, Homeownership Rates For The U.S.:1965-2007, (October 26, 2007), http://www.census.gov/hhes/www/housing/hvs/qtr307/q307tab5.html#9_8_06.pdf.
3. Richard K. Green and Susan M. Wachter, "The American Mortgage in Historical and International Context," *Journal of Economic Perspectives* 19 (2005): 100.
4. The Mortgage Lender Implode-O-Meter Homepage, <http://ml-implode.com/>.

The Mercatus Center at George Mason University is a research, education, and outreach organization that works with scholars, policy experts, and government officials to connect academic learning and real world practice.

The mission of Mercatus is to promote sound interdisciplinary research and application in the humane sciences that integrates theory and practice to produce solutions that advance in a sustainable way a free, prosperous, and civil society.

Jerry Brito is a senior research fellow with the Regulatory Studies Program at the Mercatus Center at George Mason University. His research interests include regulation, telecommunications policy, and government transparency. Mr. Brito is also a contributor to the Tech Liberation Front.