

Free Culture 2008 Conference, Politics & Transparency Panel

October 11, 2008, U.C. Berkeley

Joshua Tauberer, GovTrack.us

Policy is the front line of the Free Culture battle. I want to convince you that civic technology is key, critical. The reason is that we live in a world of increasingly increasing information. The U.S. Congress generates its fair share of data, considering some 10,000 bills each two-year session. The Farm Bill this year, which sets up agricultural subsidies and community food programs, was 650 pages long. It would take 24 hours to read through the bill --- and that's just about the time Members of Congress were given by House leadership to read it before voting on it. On the screen are earmark totals for this year for congressional leadership, with indicted "series of tubes" senator Ted Stevens securing the most earmark money. Taxpayers for Common Sense reported that last year there were 13,000 earmarks. That's a large number, and I'm not talking about the money, right? That's a large number for anyone to process: The problem is we just can't tell what's going on in Congress, what our representatives are doing, or what kinds of conflicts of interest or wasteful spending there might be. It's a weird perversion of 'security through obscurity', something like 'hegemony through verbosity'. This is what the public is up against to be involved, to get good policy made.

A very smart guy Clay Shirky at the last Web 2.0 Expo said, speaking generally, that we should start accepting information overload as a fact of life, and if we can't cope then we have to rethink our *lives*. When it comes to social networking, maybe Clay is right. When it comes to civic participation, he's dead wrong. If we don't manage the information coming out of our government, then there's no chance we're going to live in a Free Culture. There's no chance that good policy will be made, and there would be no point to a conference like this.

To have real civic engagement, there needs to be civic hacking. We need to be developing tools, new technology, to help the public cope with the information produced by our government and to help them engage with our government. But this isn't only about making a Free Culture top-down. This is about remixing government data, creating transformative views into the government.

So my claim to fame is GovTrack.us. GovTrack screen-scrapes information automatically from government websites and re-presents it in a number of new ways. The site is both a research tool and a customizable congress-tracker. On the research end, you can search and read current bills in Congress like this one, find your congressional district with a Google Maps mash-up, learn about your representative, how they voted, and see some legislative statistics. On the tracking end, you can subscribe to the thousands of feeds throughout the site, such as for bills and subject areas, and get updates by RSS or email whenever congressional action occurs in the area you are tracking. There are also widgets that you can embed on your site.

Actually the site is three things. Besides research and tracking, the site is an open database of legislative information, since our government doesn't provide one itself. You can download all of the raw data on congressional legislation, legislators, votes, and speeches, combine it with other information, and make your own new views into what Congress is up to. To give you an idea of what's possible, I'll mention a few dot-orgs that do this already. OpenCongress is kind of a GovTrack clone that is taking a more social angle, from the same people that have brought you other free culture projects. (Ah, the problems of forking.) MAPLight tracks the correlations of campaign contributions and votes at the federal level and in California --- this is showing that industry groups supporting a bill gave more to those voting in favor of a bill than to those voting against. There are some reasons

to take these numbers with a grain of salt but it's still interesting. OpenCongress and MAPLight are both funded by the incredible Sunlight Foundation. FollowTheOilMoney does something similar to MAPLight, tracking money and votes. MAPLight and FollowTheOilMoney combine vote data from GovTrack and campaign finance data from OpenSecrets.org. Larry Lessig's Change Congress website uses a map based on a GovTrack API to color code congressional districts by the amount of PAC money the incumbent takes. And there are a few others. The raw data on GovTrack is public domain.

GovTrack is open-source. The source code isn't pretty, but you can download it, under the AGPL license. Until recently the site has just been me, but in the last few months I've had a handful of people start to get a little bit involved. The site's design was recently redone by someone who volunteered out of the blue --- which was really great because I don't have any design skills. I also have someone working on a new bill text viewing system that helps particularly for these 650-page bills, so you can drill-down easily into sections, and to see side-by-side comparisons of the text of the bill as it changes through the legislative process.

The site pretty much runs itself and it's just a pet project for me. My day job, so to speak, is being a grad student.

I'm not the only person working to combat information overload in civics --- Michael Dale, who is in the audience somewhere, has a project MetaVid that is a transformation on C-SPAN video --- but the community is not very large, and it's not very cohesive yet. In fact, we're much like small islands, unfortunately. You can find some of us in my Facebook group oGosh!, that's Open Government Open Source Hacking. If you're not on Facebook, you can also check out the oGosh wiki for links to relevant mail lists.

There can never be enough people in this effort. I have a laundry list of ideas I'd love to get help with for GovTrack. But I'm going to mention a few other projects just in the area of giving the public the ability to read the law of the land --- something as basic as reading the law --- and ways those projects need help, also to give you an idea of what this community is like of hacking on civic projects like these. U.S. law has four parts. The Constitution is pretty much the only part of the U.S. law that's, say, taken care of in terms of being put online. The United States Code, which is the output of Congress, is put online by a project at Cornell. Tom Bruce who runs that says he needs some help from web developers, someone interested in text processing and NLP, or legal wonks to create secondary materials. The next part of the law is federal regulations. Would you believe that if you want an electronic copy of the Code of Federal Regulations, the Government Printing Office asks you for a check for \$17,000? Carl Malamud, the Robin Hood of political data, gathered up enough cash to purchase it so it can be openly posted online. We're getting those files now and we really need someone to step up and turn them into a user-friendly Web 2., 3., whatever-point-oh website: something better than the government has. For case law, the fourth part of U.S. law, Stuart Sierra over at AltLaw.org is creating the first free searchable and semantic database of Supreme Court and appellate decisions. He needs help too.

And that's just federal law. Other areas where technology needs to be applied to civics are the law-making process (that's my area), the regulation-making processes, state law, state legislation and regulations, local law, and voter information.

There's a lot more to be done, but I'll leave the rest to your imagination. Maybe we can hack together tomorrow during the unconference. Thank you.