Cato Gives Congress Low Grades On Open Legislative Data Transparency

September 26, 2011 by Alexander Howard

The school year may have just begun but Congress has already received an early report card on the transparency of its legislative data. The verdict? A 2.47 GPA, on average, if you don't include the 4 Incompletes. That's on average a bit better than a C+, for those who've long since forgotten how grade point averages are computer. It also means that while Congress "passed this term," any teacher's note would likely include a stern warn that when it comes to legislative transparency, the student needs to show improvement before graduation.

Rate Congress Transparency Report Card

Jim Harper, director of information policy studies at the Cato Institute, analyzed the "Publication Practices for Transparent Government" and found it a bit, well, wanting.

If you're interested in opening up the United States federal legislative system, you can tune into a livestream of special DC forum this morning where Harper and other open government stakeholders "rates Congress. Brandon Arnold, director of government affairs at the Cato Institute, will moderate a discussion between Harper, Rep. Darrell Issa, chairman of the House Committee on Oversight and Government Reform, and John Wonderlich, policy director at the Sunlight Foundation.

A better data model

The Cato paper analyzes Congressional achievement through the lens of four basic concepts in data publication: authoritative sourcing, availability, machine-discoverability, and machine-readability. "Together, these practices will allow computers to automatically generate the myriad stories that the data Congress produces has to tell," writes Harper in a blog post today. "Following these practices will allow many different users to put the data to hundreds of new uses in government oversight.

That data model used to produce this analysis should be of interest to the broader open government data community, in terms of a good matrix for rating a given legislature. "Data modeling is pretty arcane stuff, but in this model we reduced everything to 'entities,'   each having various 'properties,' explained Harper.   "The entities and their properties describe the logical relationships of things in the real world, like members of Congress, votes, bills, and so on. We also loosely defined several 'markup types' guiding how documents that come out of the legislative process should be structured and published. Then we compared the publication practices in the briefing paper to the 'entities' in the model."

While the obvious takeaway is that Congress could do better, Harper gives the Senate and House due credit and time to improve. "This stuff is tough sledding," he allowed. "The data model isn' the last word, and there are things happening varied places on and around Capitol Hill to improve things. Several pieces of the legislative process nobody has ever talked about publishing as data before, so we forgive the fact that this isn't already being done. If things haven't improved in another year, then you might start to see a little more piquant commentary."