

Digital Democracy: The World Wide Web Consortium weighs in on government transparency

September 24 2009



Image: Christine Daniloff

(PhysOrg.com) -- On May 21, the day the White House unveiled its Open Government Initiative, it also launched the website data.gov, which put information like Medicare cost reports, residential energy consumption and state-by-state toxicity reports online. Finding new technical means to make data accessible is central to the Obama administration's plans for increasing government transparency, and as those plans unfold, the administration will have growing support from the World Wide Web Consortium's eGovernment interest group.

MIT's Computer Science and [Artificial Intelligence](#) Laboratory (CSAIL)

is one of three hosts worldwide for the World Wide Web Consortium (W3C), which develops and approves the technical standards for the web: it's the organization that gave the world the markup languages html and xml, among other, more exotic standards. Two years ago, it created the eGovernment group to investigate ways in which Web technologies can give citizens greater access to government, considering questions like how to verify the identity of people accessing personal data on government websites, what media should be used to disseminate what information, and how to enable people with disabilities access to government sites.

Now, however, the group is redrawing its charter to focus on a single question: how best to put government data online. A draft of the new charter has been posted on the W3C website and is open for public comment until Monday.

"I have a great deal of respect for the W3C," says Aneesh Chopra, who was confirmed as the United States' first chief technology officer in August. "Its contributions to advance the president's Open Government Initiative are both useful and a good source of feedback to ensure we are delivering to the best of our ability on the president's vision."

Unlike most of the W3C's other groups, the eGovernment group isn't developing new standards. But according to Sandro Hawke, the group's technical lead, it is drawing on W3C's contacts in both industry and government and its expertise in facilitating conversation. "They have been wonderful about really galvanizing volunteers and creating ways to usefully organize the work," says Beth Noveck, Chopra's deputy for open government, "where their know-how about process as well as their know-how about the substantive things is very useful."

The eGovernment group has about 200 participants representing roughly a dozen U.S. federal agencies, governments in South America, Africa,

Europe, and Oceania, international development organizations, and major manufacturers of computer software and hardware. About a quarter of those participants have joined in the last week, as the group has been trying to build momentum behind its new charter.

The group signaled its change of direction in early September, however, when it posted a first draft of its guidelines for publishing government data. The draft also has an associated wiki page, where group members can propose revisions online. "In working on plans for the open-government directive," says Noveck, "I've read through their work more than once, because it's evolved and changed over time and gotten additional contributions."

Brass tacks

The document advises governmental organizations not to worry, initially, about making data pretty for online presentation but simply to post them in their raw form—preferably a "structured" form, where, at the very least, columns and rows of data are labeled. "You don't turn a spreadsheet into a PDF and then send the PDF unless you want to make sure that someone can't get at the numbers," says Hawke. "It's the kind of thing someone might do to dodge the mandate, if the mandate is to put data online."

The logic behind this and many of the group's other recommendations is that posting machine-readable data now is more useful than posting human-readable data later. That's because machine-readable data, however chaotic to the human eye, gives hundreds of millions of interested parties the chance to create programs that mine it, recombine it, and present it in any way they see fit. That, the W3C believes, will yield useful results more quickly than the labor of a small team of overworked developers trying to predict how the data will be used and designing accordingly.

Indeed, making data on the web machine-readable is the overarching goal of the W3C as a whole. The html and xml standards are well established; the W3C is working now on the standards that will define the so-called [Semantic Web](#). If the current web is like a disk full of word-processing documents-you can either summon a page by its name or search for words it contains-the Semantic Web would be like a database, where every item of information is categorized, and new queries can combine categories in any imaginable way.

The W3C has published several Semantic Web standards, and while they have a few high-profile adopters-the New York Times, for instance, has said that it will use Semantic Web technologies to organize its entire archive of articles-they are by no means widely used. Although the eGovernment group's shift of focus is too recent for it to have settled on any long-term research projects, Hawke predicts that its members will want to develop demonstrations that show how useful Semantic Web standards can be in organizing government data.

"What people have referred to a few times as the killer app for [government](#) data is just localizing everything," says Hawke. "Right now, all the data flows to Washington and then sits there." But with Semantic Web technologies, he says, local communities could build applications that automatically pull together data from the scattered web sites of different federal agencies, Congressional or Senate committees or subcommittees, and courts and "let people see anytime a regulation is coming along that is actually going to affect their neighborhood or job status."

Provided by Massachusetts Institute of Technology ([news](#) : [web](#))

transparency (2009, September 24) retrieved 25 April 2024 from
<https://phys.org/news/2009-09-digital-democracy-world-wide-web.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.