

Takeaways from the NSGIC Midyear 2011

Monday, March 7th 2011

By [Adena Schutzberg](#) - *Directions Magazine*

Summary: *Some GIS conferences are full of pearls of wisdom or big announcements or exciting new technology. Others are more thoughtful, calm, more suited to reflective thoughts and long-form journalism. This year's National States Geographic Information Council (NSGIC) Midyear, held in Annapolis, MD this past week, falls into that second category. Here then, is one takeaway in particular that spanned many presentations and topics, plus a few more focused, but equally important takeaways.*

Crowdsourcing is Here to Stay

Perhaps because of threats of someone “judging the twitter stream,” [tweets were abundant](#) (and duplicative). There's a public [Google Doc of “notes”](#) that's valuable for those interested in the “blow by blow coverage.” First, I'll discuss a thread that popped up early on the first day, and I felt, linked the presentations and conversations. That thread relates to crowdsourcing and its special case, volunteered geographic information (VGI).

Crowdsourcing turned up in the very first set of presentations that addressed the Transportation for the Nation ([TFTN](#)) effort. After an update on the writing of the strategic plan by Steve Lewis of USDOT, Michael Turner of Applied Geographics asked for input on three of the “grayer” areas of the work:

1. Can federal agencies collaborate at that level (USDOT and Census, for example)? If not, why not?
2. What's the role of VGI/OpenStreetMap (OSM) and their relationship to “authoritative sources”? OSM is considered a source for TFTN in the current thinking.
3. What would a public/private partnership model look like? What is its potential? Can state models scale?

The feedback on the second topic was positive, with one commenter pointing to the success of the broadband mapping effort and its use of crowdsourcing (more on that later). This tweet appeared: “Open Question #nsgic: what states agree that OSM is a ‘reliable’ data source?” I saw one reply: “osm has issues, but in general, crowd sourced data does fit somewhere in the bigger picture...a good feedback loop key #nsgic.”

Discussions followed concerning the National Enhanced Elevation Assessment, the NSGIC state summaries and [Version 4 of the GIS Inventory](#), and all highlighted the fact that the data in those projects were in large part volunteered, and are, I would add, pretty complete. The number of people and datasets participating speaks volumes.

Dr. Raphael Bostic, assistant secretary of the Department of Housing and Community Development (HUD), didn't address crowdsourcing directly, but he noted what might be considered the reverse of crowdsourcing: seeding the landscape with geographic knowledge and training. In particular, he noted HUD was looking to partner with the University Consortium for

Geographic Information Science (UCGIS) to form a sort of GIS extension service, modeled after the Agricultural Extension Service.

After a discussion of the Geospatial Multistate Archive and Preservation Partnership ([GeoMAPP](#)), representatives from one state that was not chosen to participate in the initial project learned that any interested state could now [participate](#).

The last presentation of the morning was from Learnon Dalby of the Arkansas Geographic Information Office. He explained that his office was tasked with “mapping the state budget” for the governor. This request came *after* the governor’s office was presented with a map showing how the governor won every county in the state in the recent elections. That sounds like a governor who “gets it.”

While there were many takeaways from Dalby’s presentation about the process and product (keep it simple, use color consistently, present data in a familiar form, i.e. spreadsheets; slides are [here](#)), one comment struck me. It was about the use of [Google Refine](#). Per Google, it’s “a power tool for working with messy data, cleaning it up, transforming it from one format into another, extending it with web services, and linking it to databases like Freebase.” Why is that important? “Because budget data is a bunch of \$)%W.”

That assessment of budget data from Dalby sounds like how many people feel about crowdsourced data. If we can clean up all sorts of tabular and geodata collected in other ways, why can’t we clean up crowdsourced data? Why is cleaning up these data a new idea? Is it time to suggest that crowdsourced data are just data? Some have argued for years that spatial data are not special. Maybe crowdsourced data are not, either.

National Telecommunications and Information Administration (NTIA) program manager of the State Broadband Data Development, Anne Neville, highlighted the process used to capture the data: having the states take the grants and develop partnerships. She acknowledged the challenges with crowdsourced data: “There are uninformed opinions,” she agreed, “but ultimately that will make this a better dataset.” The next data update is April 1. And, it will include the crowdsourced feedback. How? The 30,000 individual pieces of feedback to the [national map](#) will be sent back to the states, which will, in turn, review the input and make any necessary changes. She concluded that the use of the crowdsourced data ensures this is an “ongoing effort to show this is a useful dataset that has policy implications.”

FCC GIO Michael Byrne dove into a technical discussion of the effort, which frankly went over many heads. Worthy of note: the implementation was fully RESTFUL and not hosted in the cloud. His lessons learned:

- Rethink the goal of 100% data, 100% functionality on launch day
- Don’t have crises just before launch
- Scale approach to your visits

Attendees heaped on lots of praise and had many questions about the project, but Drew Rifkin of Safe Software summed up the success of the national broadband effort better than I can: The

effort created the first layer of the NSDI. And, I'll note, crowdsourcing is a key part.

The Corporate Leadership Council session gave sponsors a chance to address the attendees. Two of those presentations related to crowdsourcing. Antony Pegg of MapQuest focused on the organization's "[open](#)" initiatives, which based on the tweets and comments, were new to many. I want to highlight two key crowdsourcing tools. One tool allows those who don't want to update geodata in OSM directly to tag errors for others to fix. That may not be the traditional picture many have regarding crowdsourcing geospatial data, but it's a valid and valuable one. The second tool is a notification that will, soon, allow anyone interested to monitor changes to a specific geography via daily, weekly or monthly alerts. My point here is to highlight that crowdsourcing and VGI are not, and should not be, limited to just raw data collection. Instead they can and should be broadened to contribute to workflows to enhance that process and to ease confirmation/correction as these examples suggest.

Jaymes Pardue of TomTom showed a map that looked very familiar. It was basically the well-known "Earth at Night" image. The source? Traces from TomTom users who had agreed to share their paths with the company. The percentage of TomTom device owners who opt in? About 90%. That's a significant validation of the possible participation and value of this type of passively collected VGI.

My sense is the NSGIC community sees the value and challenges of crowdsourced data and is convinced for the most part that the former outweighs the latter.

Laws Force Data Duplication

Bill Johnson from New York gave a great example of how current U.S. laws force data duplication and thus payment duplication. Since his state (and all of them, actually) cannot by law use the point GPS address data collected by Census, New York State got a grant from NTIA to produce the same point data. It's time, many in attendance suggested, to update the laws that force this sort of duplication.

Three National Vertical Data Collection Efforts that Worked (with credit to Bert Granberg, Utah)

The [National Elevation Dataset](#), [EPA's Environmental Information Exchange Network](#) and the recent [broadband mapping](#) efforts are examples of different ways to capture national datasets. Each has been successful in gathering good, mostly complete data and serving their constituents. While it's unlikely these models can be mixed or matched for any purpose, they are a great starting point as we move forward with our various nationwide efforts, including [geoplatform.gov](#).

Addressing Addresses

Many agencies need addresses (Census, U.S. Postal Service), as do local responders to E-911 and NG911 calls. But, the needs/goals are different and there are legal barriers to data sharing. All that said, the members on a panel on the topic (public, private, federal, local) agreed we need and want better addresses. NSGIC's own Bill Burgess asked: Is it possible, as was done for Transportation for the Nation, to have an agency at least begin to explore a united national addressing effort?

Another Organization to Advocate for Geospatial: NAPSG

The National Alliance for Public Safety GIS ([NAPSG](http://www.napsgfoundation.org/)) “was formed in 2005 to overcome the challenges faced by Federal, tribal, state, and local public safety agencies. It was reorganized in 2008 to operationalize the mission through education and research around information sharing and data interoperability associated with GIS and advanced technologies used by the public safety and homeland security communities.”

<http://www.napsgfoundation.org/>

Other Efforts to Keep in Mind

Census 2010 data are being delivered. Redistricting is coming. NG911 efforts are underway. All of these have implications from the local to national level. The fourth version of the geoplatform.gov document is at the Office of Management and Budget and should be public soon. Public safety and homeland security still demand our attention. This NSGIC meeting, like all others I’ve attended, highlighted the value of relationships between the various agencies and levels of government, and of continuing to strive for the best data possible to support all of the public sector’s work.

NSGIC 2011 MidYear Meeting Notes – Annapolis, MD

[Directory View](#) of notes and other items from this event that are held in Google Documents.

The Back Channel:

- [#NSGIC-tagged tweets](#) -- on Twitter
- [Spreadsheet of #NSGIC-tagged tweets](#) -- Captured using TwapperKeeper. From 2/27/11 (04:24:21) through 3/2/11 (23:12:16) Hopefully we got them all
- [Word Cloud of #NSGIC tweets with RT'd profile names left in](#) (LearnDalby was most-retweeted)
- [Word Cloud of #NSGIC tweets with RT'd profile names removed](#)
- [Links Tweeted from NSGIC MidYear](#) (Not including the [deck.ly](#) links that allow for longer tweets)

[NSGIC Glossary](#) -- This is a working list of acronyms that cropped up at this conference. We're crowd-complaining about them and crowd-sourcing the definitions. If you have an acronym that stumps you, add it and leave the definition blank. If you can fill-in any blanks, please do!

Session Notes -- Rough notes taken when we were in the room and paying attention. There are gaps where we had to take a bio-break, get more coffee, or make our own presentations.

Monday, 2/28/11

- [Notes by Joy Paulus](#)
- Notes by Mike Mahaffie
 - [2/28/11 Morning Session](#)
 - [2/28/11 Afternoon Session](#)

Tuesday, 3/1/11

- [Notes by Joy Paulus](#)
- Notes by Mike Mahaffie
 - [3/1/11 Morning Session](#)
 - [3/1/11 Afternoon Session](#)
- Breakout Sessions
 - [Census Bureau](#) -- Notes by Mike Mahaffie
 - Geospatial Platform
 - NTIA/FCC
 - [Notes by Joy Paulus](#)
 - [Notes by Joe Sewash](#)

Presentations are posted in various places. Eventually, all of the presentations that can be posted will be posted on the NSGIC web site.

- [Strategic Planning for Transportation for the Nation \(TFTN\)](#) -- Steve Lewis, GIO, USDOT
- [Putting a State Budget on a Map](#) -- Learn Dalby, State of Arkansas
- [Communications Committee Report](#) -- Mike Mahaffie, Delaware