Session 2: Wednesday 5/6/2015, 3:15pm-4:15pm

Location: Pittsburgh/Erie

Session Title: Regional Data Centers

Organizer: Scott Gaul

Primary Notetaker: Rob Pitingolo

Participants: Bray, Tatian, Pierson, Gaul, SIlver, McKieren, Young, Sanders, Simmons, Kim, Lee Hochleitner, Flodine, Cowen, Morehead, Pittman, Capperis, Urban, Forrester, Manieri, Gradeck (late)

Notes:

Gaul: I proposed this session but don’t have a ton to contribute. Hartford is a tiny city but in a big metro area. I’ve heard that some cities/ metros have regional systems. Region is largely suburban and there’s a bright spot for open data in the city but in small cities and towns it’s a bigger challenge. Anyone from Pittsburgh?

Forrester: I’m from Allegheny County.

Gaul: CT abolished counties which is what leads to all these little municipalities. Anyone cover something beyond a single town.

[Hands raise]

Simmons: We have Mecklenberg county and then also six towns as part of that. We get police data and some other small things. City of Charlotte gives us the most. University Area is a 14 county region including rural areas that doesn’t give us much

O’Brien: In Milwaukee we have a mix of city and county data. It’s a hybrid.

Bray: We polled police departments about getting crime from all of Dallas county. It was an unhealthy experience. Dallas (city) has open data so it’s there already. But we had large police departments who said no, it’s not public information. Some small departments printed PDFs of their crime reports. They were going to fax them to us. It’s shocking because people read the UCR and assume it’s all available. There is no incident level shared data system. They don’t talk to each other.

O’Brien: Anyone know of the name standards?

Bray: There is an xml standard for all of this data. Part of it is in the NIBERS platform. Local agencies in Dallas won’t let you into their RMS. It’s a conscious decision to not make data available.

Flodine: When we built an open data platform we had a hard time getting people to agree. We published a lot of the same data, just in machine readable format.

Gaul: Could be easier to respond to requests if there's a common person staffing the data organization.

Manieri: Wouldn’t you still need agreements (legally) from each municipality? Even if the data was collected by a single person?

McKeiren: We serve a 12 county area, mostly rural. When we go outside the city it’s been very spotty. Haven’t really had much need to try to get standardized information. Outside San Antonio the counties hold just about everything we care about.

Cowan: There are hundreds of municipalities outside Chicago. Consent is a big issue. Cities are used to getting requests for data. Small suburbs don’t even have a staff necessarily. There’s a lot of good data at those level nonetheless. If I want court data, I have to go case by case. So consistency is an issue. Municipal data is erratic. Funding is a big issue.

Tatian: There’s no economies of scale. 2 places, 2x as much work. 5 places, 5x as much work.

Cowan: If you standardized it, there could be economies of scale.

Gaul: Every other town in our region is a customer of ESRI. Could use that kind of thing to build an economy of scale. Are there economies of scale in specific areas? The other question is about consistency. The law is clear about what you have to disclose. I’ve heard Pittsburgh has made some progress here. Is there anywhere else?

Bray: We find discrepancies because of how data was collected and coded. There’s a policy layer that just doesn’t work. Grade inflation a good example. C grade is not a C everywhere. Look at hospital data and try to figure out why prices are different. It’s hard if you think the procedures aren’t actually the same. When places don’t want to be compared, they either tweak the data or shut it down entirely (no data sharing). Open data is forcing the hand. The conversations are accelerated by the technology.

Gaul: What’s the role of NNIP is this?

Manieri: Sounds like advocacy is what NNIP could bring to the table. Why open data and such is necessary. Funding would probably come from local funders, who aren’t usually regional. Could advocate for more data for comparison purposes. Anyone else live or work in a region where we’re struggling to even define a region? Don’t want to compare Newark to NYC. Maybe not to less urban parts of NJ.

McKieren: More about a “complete picture” rather than a comparison perspective. Oh hey, there’s Bob [points to Bob Gradeck].

Gradeck: Thought I’d stop in. We persuaded people by talking about how we could solve problems better collaboratively. We talked to the mayor’s office and they were able to pull in the county executive. They walked into a foundation and they opened the checkbook. Mellon is the first funder and Heinz. UPitt has been providing tremendous support. They provided a data sciency person. We started with what’s important to funders and what’s important to our funders. Next big issue I think is stormwater data.

Manieri: Your foundations care about regions?

Gradeck: They focus on local work. It’s kind of an anomaly the amount of money we have per capita. A lot of development downtown because of them.

Baldwin: What’s the relationship between Pittsburgh and Allegheny?

Gradeck: City in the county. County does health, courts. City does housing, property tax, garbage, police, fire.

Gaul: How about sustainability?

Gradeck: We made sustainability a priority. We host everything with UPItt. We invested in UPitt instead of other stuff. We have a programmer and someone doing project management.

Gaul: Stormwater is so unsexy. That’s interesting.

Gradeck: If their data is important based on how it can be used. We will help automate publishing, standards. We acknowledge that we don’t know everything but try to bring everyone into a collaborative process. And we’re just making stuff up. [chuckles]

Parks; What models have been useful?

Gradeck: Open data Philly. Chicago has been in this for a while. Don’t invest in technology, invest in people.

Gaul: Anyone else work closely with a Library system? Becca, would you say they gave people stuff to think about?

Lee: They were awesome.

Gradeck: They have an office equipped with people to help with data management. They are there for support with other projects. Even people like Greg [Sanders] have been helpful. Just talking to people like Greg.

Sanders: We’re all proud of you Bob. People are impressed with what Pittsburgh has done.

Gaul: One question Peter raised is about incentives. The reason this exists here [in Pittsburgh] is for this.

Gradeck: I didn’t have to talk about data. It was the other people who sold it. They ought to tell the story. I can’t tell their story. You can start small. Albuquerque started small. See what the reaction is and then scale it up. You can use a tool that pulls csv files out of any website. We have a page on the NNIP website that describes our experience on this. I would like to know. Where is everybody else?

Bray: In awe.

Gradeck: We haven’t done anything yet!

Bray: Which is still more than most.

McKierien: We have an FTP in Jinga. It’s a start.

Gradeck: Open Data portal stuff has gone a long way. Some is pretty cheap.

Cowan: We have a six county portal. We now host the CCAN data site. Take a lot of effort to keep up. We struggle with funding.

Gradeck: Who publishes?

Cowan: Most of the data we get is free. We purchase two sets. Neither has a problem with us presenting the data we do get at the tract level as long as it’s not the parcel level data.

Flodine: Where do you get income data from?

Cowan: Easy analytics from NY. Natoinal data.

Gaul: What’s going well outside of Pittsburgh. SA example is interesting.

Pitingolo: Time’s up.

[end of session]