Session 1: Thursday, September 15th 3-4pm

Location: Stouffer

Session Title: Data Collaboratives

Organizer: Raleigh

Primary Notetaker: Pitingolo

Attendance: Raleigh, Morgan, Morris, Kingsley, Johanson, Linn, Abraham, Grossman, Pittman, Urban (Noah), Ozuna, Spike, Rouault, McKieren, Gradeck, Urban (April)

[Start 3:05]

Morris: I’m challenged as a data manager when I used to do GIS and I’m good with community engagement.

Kingsley: I’m curious how cities think about the data tables based on the issues that form. My sense is that cities need integrators.

Johanson: Curious to learn about how people have been able to integrate data from different entities. We have multiple partners interested in the same thing I’m curious how we get them to be more likely to share.

Spike: We’ve have lots of collaboratives and they’ve all been weak.

Linn: We’ve done collaborative projects before. We’re trying to define and gather violence data in the city. In terms of data collaborative. So many start differently and process for sustaining I’d be interested.

Abraham: Started as the regional data collaborative. There are 250 people we work with that have data. There’s no one one pull them all together. Having 2 or 3 staff to pull is impossible. We have 2 million hospital records but can only use it for one purpose.

Grossman: Everything we do is a partnership. Talking about the process of coming to a shared goal is something. THe other is the distinction between a collaborative project and an ongoing collaboration. What can Microsoft do to make that happen.

Pittman: I want to clarify we’re not talking about an IDS? Just sharing data and being able to. OK good. South FLorida council wants to do a data commons. The trust just does children and families but there’s so much more. I’d love to partner with universities and government to see through that on a regional basis.

Rouault: We have a data common, built 5-6 years ago in states of disrepair. I’m responsible for rebuilding it. I’d love to get best practices and relationships with the kinds of organizations that contribute. We collect from 101 cities around Boston. We want to know how to make it valuable and not replicate what the state does. Before that I was running a startup that supported collection of place based data. A lot of that turned into data commons.

Urban: In the weeds to implement a formal collaborative in Detroit from Motor City Mapping. I’m here to learn and take in and hopefully take that back and get the money we need.

Jordan-Detamore: Previously of Prov Plan. Looking to get the same data sets from adjacent jurisdictions is a big issue.

Raleigh: I nominated this session hoping to learn, now I’m facilitating it.

Morgan: We helped Pitt launched the western PA data center. We’re mostly as a data publisher. We’re kind of a data intermediary with all those other counties too.

Ozuna: We are embarking on creating a data collaborative. I wanted to hear about experiences.

Richardson: I’m interested in sustainability of these collaboratives.

Raleigh: OK, let’s toss some things out. There’s the integration piece. There’s what’s the value added or the incentive. What’s the common interests. Anyone seen these things come together?

Abraham: When there’s a major grant due everyone contributes.

Kingsley: Co-sign. Sometimes foundation funded efforts make no connections to data resources. My question is why they’re not aware of resources or why it’s not for them. Maybe jurisdictional issues.

Raleigh: Is it jurisdictional? Trust? Lack of awareness.

Morgan: Bob talks about the large property assessment file which was the first we published. It was really intimidating to community groups. We didn’t know why it wasn’t changing their work until we had the user groups and build an extractor.

Rouault: Our most popular feature is a community profile. There’s also a moment when they data they want for their story doesn’t exist.

Kingsley: Often people want a cut that’s not available through the tool. It’s not always clear who to call or whether you can use their time.

Richardson: In mke a group does after school activities. Data literacy is a rainbow. We can’t assume that people know what question to ask. If they have the data, what to do with it? The question had to be - what to do today? What to do now? Have to understand where people and then meeting them there.

Raleigh: We’ve identified different groups to target. I’m a data producer. There’s the public at large. There’s power users, journalists, etc. There’s how to get it into hands and how to get people to put things in. Sounds like Mark has health data. Do you have a lot of people contributing and how?

Abraham: Universal to our work, there isn’t funding or resources to push data out. It’s time intensive as you know.

Richardson: Crowdfunding. Get a bunch of people involved and then have them pay?

Abraham: Foundations will do that. Each put in some money.

Kingsley: I figure cost is less an issue than trust. A lot of admin data is crap. I haven’t seen it done outside of specific initiatives like collective impact. Even then it’s hard to keep it for more than a year.

Rouault: We launched a repository for the state for planners or GIS people who are asked to contribute data. There’s not really a value proposition. One thing to do is use it as a search tool. So use technology to help governments then they pay us in data.

Morris: Joy made that point. Show the tool that’s great but you have to get your data into that format.

Richardson: I’m calling to ask about a neighborhood. Maybe you don’t know where the info is. I might be angry. But there’s a pain point there. Anyone had success finding that point and squeezing it harder?

Raleigh: That worked in Detroit to get stuff into the portal. Low hanging fruit is done. Now it’s about getting hard stuff in there. One way of selling it is the number of FOIA requests they spend time on over and over that they wouldn’t have to do anymore. “Thanks for calling - go here”.

Gradeck: Is this a session where I walk into an ambush?

Morgan: We’re trying to get municipalities to be data publishers. We’re trying to get the state crash data for our county. We mapped those things and did cluster analysis. Who owns the roads that these crashes occur on? We see this thing in your municipality so let’s talk about that. Or maybe it’s owned by someone else, so figure out how to add to the analysis. I only have one code enforcement person who is swamped.

Gradeck: It’s all about however you can do it. We’re looking at other issues and trying to figure this out. The county would have people to take the file, review, document, then work with IT to put it onto a ftp server. Then we pull it in and make sure it’s all running properly. We can provide that technical capacity for all kinds of folks. It’s about making it work in a distributed way.

Kingsley: Data is acquired over time. There’s a project or grant that starts it. How do you think of sustaining it and institutionalizing it? So you don’t just lose a bunch of data every year.

Raleigh: Good question.

Morgan: We’ve seen in the county that how partners share with each other is manual and it breaks down because of the people involved. It’s not the database that’s changing but the people running the reports. We’re trying to build infrastructure to keep it going.

Rouault: Providing value on top of data is important. It needs to have a user for the future. When the grant money dries up you have to move on.

Gradeck: If number of data sets is important, that’s not a good measure. Gotta find the value. We had someone from the CDC come in. Maybe we don’t do it, but it’s a student project or a brigade project. We try to find a marketplace for this. People do things and don’t always know we can help.

Grossman: People with skills to do all this stuff is hard to find.

Gradeck: We’ve got data science students are Carnegie Mellon. It’s building their skills, they go off and then help us out. We don’t know if it will work.

Grossman: Training people with those skills.

Richardson: What do you mean?

Grossman: Microsoft has a week hackathon. Some are internal, some are external. One in Chicago was with a housing center. The reason it went well because we took someone who does enterprise customers and helped this housing center. I can design a tech solution for anything but it’s not usable.

Spike: We saw when you take out data geeks from decision making tables and the people left don’t know shit about data holdings and they’re making bold declarations about what they need, we’ve got a school superintendent with measures that aren’t measurable. Otherwise you end up with things that aren’t feasible. Then it’s embarrassing when these people are on the hook for this.

Richardson: Lean process. Human centered design requires all stakeholders in the room. I adopt that as a business practice and will bring back to NNIP. Are people actively using that?

McKieren: People don’t have time to come but have time to complain later. You can’t be psychic and design something cool that doesn’t meet any needs. We get paid to sit in rooms and think things through. We hired for this ability. Someone will come and say we need you to do big data and we have $3k.

Richardson: Step one is discovery and costs $3k.

Rouault: Who pays for human centered design?

Gradeck: We’re going to use some money have consultants figure this out, then open source that. If we can speak that language and show what we did, maybe you can run the workshop.

Grossman: Is this an allowable or unallowable expense?

Spike: With open data you need a local example. You can’t give a Boston example in Oakland they don’t give a f\*\*\*. They have to get that it’s not just for other people. It’s so important for us to write about this stuff. Otherwise you think dammit I knew April did that thing but I can’t remember what.

Kingsley: Do we all share a definition for a data collaborative? Do any of the funders understand that?

Raleigh: I like being able to say no. With a collaborative sometimes the topic is too important not to get together on, like Motor CIty Mapping. For us that’s the one time it actually worked. The political willingness and resources coming together all made that happen. We’ve got 10 minutes left.

Richardson: Has everyone done business model canvas? Does everyone have a multi sided model for customers. Facebook has multisided, it has users who don’t pay, but need users to manipulate to make money somewhere else. The first customer is the data stewards. Who are those customers? Do we treat them like customers. Now we’ve got data, now what to do with it to create value to others or back to them and provide that data as a service.

Spike: We’ve tried to make that part of the pitch.

Raleigh: I do wonder and think it’s a struggle we have. Facebook gets to generate revenue on ads and so on. What do we do when the mission is about transparency of data and we have a commitment to low income communities who can’t pay for data. What we do for food money is other value added on top of that but data are always free.

Kingsley: A partial answer is foundations. Look at social impact financing.

Raleigh: I heard about building in a tax line for public utility.

Gradeck: We have a regional asset tax. I’m trying to figure out how to do this after the grant runs out but hopefully we’re indispensable.

McKieren: Build in a margin for all of your fees, 5-10% more, then you build up a pile of discretionary funds to spend on things other people don’t see value in yet. We’re working actively to transition to build that in.

Morgan: I want to pull in the guy from the cle foundation yesterday. Places are going to have to see us as indispensable.

Raleigh: OK, who has a good closing comment.

Pittman: It would be nice if we had a definition of data collaborative.

Spike: Sharing data between more than 1 org.

Gradeck: Throw in the word infrastructure.

KIngsley: And storytelling.

Raleigh: I always here Strive, but that’s all we’ve got.

Spike: If you don’t have millions of dollars the model doesn’t work. Angela Blackwell pulled it out a little. But there aren’t many in the country. It’s lifted up as a viable that isn’t attainable. What’s the point. If you can’t scale a model to 100 cities it’s a BS model.

McKieren: No one knows data intermediary, but they get backbone if you say it like that.

Spike: We’re trying to rethink that concept. Backbone is super uncool. I don’t want to make my career being a backbone. Calling it that suggests you can forget about it.

Richardson: I’m more positive on Strive but they’re working on a definition for their model. It’s uncertain how that will fall.

Rouault: Shape shifting philanthropic lexicon. Need a glossary of terms.

Raleigh: OK 11 minute break til the next one. Thank you.

[End 4:04]