District of Columbia Interagency Data Team

August 9, 2017
2:00 PM – 4:00 PM
Office of Chief Technology Officer
200 I Street SE, Washington DC, 20003
Conference Room 1001A & 1001B

Agenda

- ***Welcome, News & Updates***
  Barney Krucoff, Chief Data Officer, Office of the Chief Technology Officer

- ***Role of the DC State Data Center***
  Charlie Richmond and Travis Pate, Office of Planning

- ***Enterprise Dataset Inventory, Lessons from Pilot Agencies***
  Panel Discussion OCTO Data Team and Agency Staff

- ***Enterprise Dataset Inventory, A General Counsel's Perspective***
  Pam Brown and Turna Lewis, OCTO's General Counsel's Office

- ***Open Discussion***

Meeting Notes
Capturing notable comments and questions from attendees.

Presenter: Barney Krucoff – House News
- Run through new highlighted data such as address apartment and condo units. DC is one of the first to publish crash data along with related tables. DC government PIPELINE contracting vehicle also published (means for DC government to hire contractors). See presentation for full list.
- Run through recent data requests from the http://data.dc.gov (public wish list). Highlighted that we do not have historic data layers. Frequent requests we share history of data. Also asked that agencies review these layers.
    - Travis from Office of Planning can help with some of the income data.

- One pager quick facts about DC.
  - This should be posted on http://data.dc.gov
- Show and tell for all the current data products from state data center. All at https://planning.dc.gov/page/dc-data
  - American community survey estimates
  - Population estimates
  - District Quickfacts
- Displayed a variety of reports and infographics of in depth topics. See presentation slides for these.
- We do forecasting for population, jobs and others over a 30 year period.
- Range of data
  - Usual demographics like race
  - Newer ones are computer and internet use and housing characteristics. Also family and living arrangements.
  - We do produce and have some employment data
  - And best of all being health insurance availability.
- Access data from the office of planning: infographics. Most are in pdf or excel sheets. This includes metadata, explanations.
- Run through of how the state data center gets data from Census Bureau. Look for "Tools and Apps" on the census bureau website.
- American Community Survey – about 10k residents get surveyed every year. Via online, paper or in person. This is what really powers forecasts in addition to the centennial census.
- American Fact Finder is the clearinghouse of all census data. Not very user friendly but good for the data scientist/analyst. State data center can help.
- There is an economic census that happens every five years – demographics of business owners, published as establishments. Includes employment, payroll, sales, industry types

- Question: what is your advice for who we should go to first? Office of Planning's State Data Center versus Census Bureau?
  - Come to us first because we likely have it in house. And if we do not then we usually will work to get it for you.

- Question: really interested in the geography changes of the census tracts. Interesting to see the progress. Is there any history of changes to lineage? Do we have the ability to relate for example, new smaller tracts that came from a specific larger census tract. Learn which tracts came from others.
  - It's a little difficult since the geographies do not nest within the larger historic geographies

- Question: can you tell us about how you use or what you can do what the American community survey?
  - We have to be careful of what geographies you are using because it is just a survey. Margin of error can be higher. Combining or aggregating with other data like census tracts can improve analysis.
  - It's not advisable to compare data that are overlapping times within those five years.

- Question: what is the methodology of the American community survey?
  - It's always 10k for the whole year but the census collects these every month throughout the year.

- Question: can you talk about the infographics that you create? Are you useing artistic software or business intelligence tools.
  - They are mostly graphic software. We use adobe illustrator and indesign. We are looking to create dashboards with tableau now that DC gov has an enterprise server for it.

- Question: as a state data center do you have access to data the census does not provide to the public?
  - We get everything the public does. However there are some datasets you have to pay for but again these are also available to the public. We do get sneak peaks of datasets prior to releasing to the public.

Presenter: Panel Discussion – Enterprise Dataset Inventory, Lessons from Pilot Agencies

Agencies (OCTO Liaison Panel Member)

- Office of the Chief Technology Officer (Suman Shukla)
- Deputy Mayor for Planning and Economic Development (Alexandre Santos)
- Department of Small and Local Business Development (Eva Stern)
- Department of Behavioral Health (Bircan Tekin)
- Office of Victim Services and Justice Grants (Dave Cook)

Questions Addressed

1. Tell us about your agency that you worked with. For example, their mission and size
2. How did you launch your EDI effort within your agency?
   a. Was it formal or informal?
   b. Any tips for what worked with your launch?
3. Were there any tricks or tips that you can provide in helping with the identification of "enterprise" datasets?

4. Which data classification (level 0 to level 4) did you find the most difficult to categorize and why?
5. How did you enter your data? Directly or did you create a working document ahead of time? As a group or some other method?
6. Did you identify a dataset that is not currently open but should be (level 0)?
7. What were the most difficult datasets to get done and why?
8. What was more difficult in the whole process:
   a. Determining your universe of enterprise datasets?
   b. Classifying your datasets?
   c. Identifying and entering your related enterprise applications?
9. Were all the fields being asked for in the data inventory tab understandable? If not, what could be better defined?

Results from this discussion helped the OCTO liaisons draft help documentation linked below. Note that these are internal documents within the DC Intranet.

[Quickstart Guide](#)

[Get Organized Prep Sheet](#)

[Frequently Asked Questions](#)

Presenter: Pam Brown – Enterprise Dataset Inventory, General Counsel's Perspective

- This was an informal discussion from the Office of the Chief Technology Officer's general counsel.
- Take away, for ADOs, include,
  - You must write clear descriptions for us to approve. We are not the experts in your data.
  - The deadline in November is for publication. Agency general counsels will need at least a month to review. So, your deadline should be end of October. For example, OCTO has 80 plus datasets for me to review!