

## District of Columbia Interagency Data Team

January 23, 2019

2:00 PM – 3:30 PM

Office of Chief Technology Officer  
200 I Street SE, Washington DC, 20003  
Conference Room 1001A & 1001B

### Agenda

- **Welcome, News & Updates**  
Michael Bentivegna, Data Visualization and Analysis PM, Office of the Chief Technology Officer
- **New GIS Database and Affects to Existing Applications**  
Mario Field, Data Curation Program Manager, Office of the Chief Technology Officer
- **Enterprise Data Inventory in Progress**  
Eva Reid, Data Visualization and Analysis Senior Analyst, Office of the Chief Technology Officer
- **Data Stewardship & Standardizing IT Contract Language for Data Collection**  
James Graham, GIS and Applications Manager, District Department of Transportation

### Meeting Notes

Best attempt to capture notable comments and questions from attendees (paraphrased). For complete presentation visit, <https://octo.dc.gov/node/1310836>

Presenter: Michael Bentivenga (OCTO) – Welcome and News

- Tableau server architecture review session,
  - OCTO provided Tableau with 3 months of data to help DC Gov plan for future use
  - Session took place on 1/22/2019
  - Our system is sized appropriately, and we have room for growth.
  - We have to do some cleanup of unused workbooks
  - We are using a good amount of extracted data which helps with performance even though takes up more space.
  - Although, we aren't using as much incremental extracts – agencies with a lot of extracts recreating the data each time

- We've found that we have a 7am bottle neck. There are too many scheduled updates at 7am. Silver lining, it's good that we have monitoring systems in place but we'll need to evaluate and spread out update time.
- We will be reaching out to agencies that tended to have space for improvements
- Tableau DC Public
  - We have some of our first sites up and shared publicly – Circulator, Vision Zero
  - This is meant for agencies to publish content that is **ready** for the public – quick load dashboards.
  - It is not meant for large internal operations dashboards.
- MicroStrategy new public environment expected launch end of February.
  - Created public facing MicroStrategy web servers, testing done and back end integration is underway
  - MicroStrategy integration guidelines for embedding analytics in Agency websites and SharePoint developed.
- GIS Software
  - Currently at ArcGIS Server 10.6
  - Coming up with 10.6.1 that will come with some streaming services
    - ESRI GeoEvent will be in production @10.6.1
    - Apache Kafka® distributed streaming platform – real time IOT Streaming data (FY 19 Q2 production)
  - *Question:* could we beta test a streaming service without letting out to the public
    - Yes, there are internal uses such as those with HSEMA.
    - Not everything streaming will have to be public.

### Presenter: Mario Field (OCTO) – New GIS Database & Effects on Existing Apps

- OCTO is migrating both DATA WAREHOUSE and GIS databases
- Applications will be most affected by the GIS database migration
- It is built on Oracle 12c – latest version
- We now have a better environment with DEVELOPMENT, TEST and PRODUCTION set up
- Setting up FME infrastructure which allows for easier transferring of data and replication
- PRODUCTION will be running on fastest Grid, hardware and Oracle database ever setup here at OCTO
- Working to have an operational production database by 3/31/2019 at which point clock starts for OCTO and DC agencies to move off of DCGISPRD
- We will be standardizing our editing workflow and documenting OCTO and agency edits
- Within DEVELOPMENT there will be workspace for agencies to edit their own table spaces where OCTO can then replicate to PRODUCTION.
- What makes new GIS database (2.0) great,
  - Current database has 16 years of various projects and datasets of which none have kept to a standardization.

- There will be better, enforced, consistency with attribution, naming and metadata
- *Question:* will tax and record lot information be included in this migration?
  - Some agency datasets such as tax lots and other major datasets will have to be offered as is. Agencies like DCRA are stewards and OCTO probably cannot enforce changes to their systems.
- We will be creating a space for archived datasets for better identification of historic data.
- Naming for web map services will stay the same but we will have to do some heavy testing to make sure switch over will not affect cloud services like ArcGIS Online and Open Data DC
- We will be able to consolidate layers such as those that have monthly and yearly breakdowns seen on open data.
- Alleviates partitioning layers published on open data and within GIS database
- Open data handbook
  - Used to be called the *data submission guide* but we scaled this back
  - This is first draft, it will grow as we continue with the data submission process
  - Includes a complete set of data set rules: naming, attribution, metadata, data dictionary.

#### Presenter: Eva Stern (OCTO) – Enterprise Data Inventory Progress & Reminder

- A run down of helpful content that is on <https://octo.in.dc.gov>
- Agency Data Officers and designees should be using <https://ediapp.octo.dc.gov>
- Asking that agencies identify at least one primary ADO even if there will be multiple editors in application
- *Question:* will open data handbook style guide be published before Feb 11?
  - Next thing to go up will be an FAQ and also list connecting ADO will Liaison
  - But the open data handbook is more detailed and pertains to publishing datasets. The EDI attribution is minimal.

#### Presenter: James Graham (DDOT) – Data Stewardship & Standardizing IT Contract Language for Data Collection

- Not something as cool as a data visualization but I would argue this is something just as important.
- Standardizing data collections is a common struggle we have within DDOT
- Here is a scenario:
  - You're a contractor and you've just won a contract to count cars and pedestrians for DC government
  - DC government says, "here is a guide of something that someone else did – just do something like that."
  - So many problems and questions
    - well how's the format coming back?
    - what kinds of info... just counts?
    - ... and so on

- DDOT is creating IT contract language for data collection
- A guide that can almost be copy/pasted into a new contract
- However we would like for other DC agencies to give us input; we can do this together
- Items such as: uptime availability, disaster recovery, minimum data loss standards, ability to reference other datasets at other DDOT teams and organizations like OCTO
- DDOT's endgame – when we have people do projects (stuff) that's for data then contractors should have this guide to reference
- *Comment:* one of the challenges we deal with is to require contractors to provide an API. Give preference points for those that can provide.

Meeting adjourned

Next meeting tentatively scheduled for March 2019