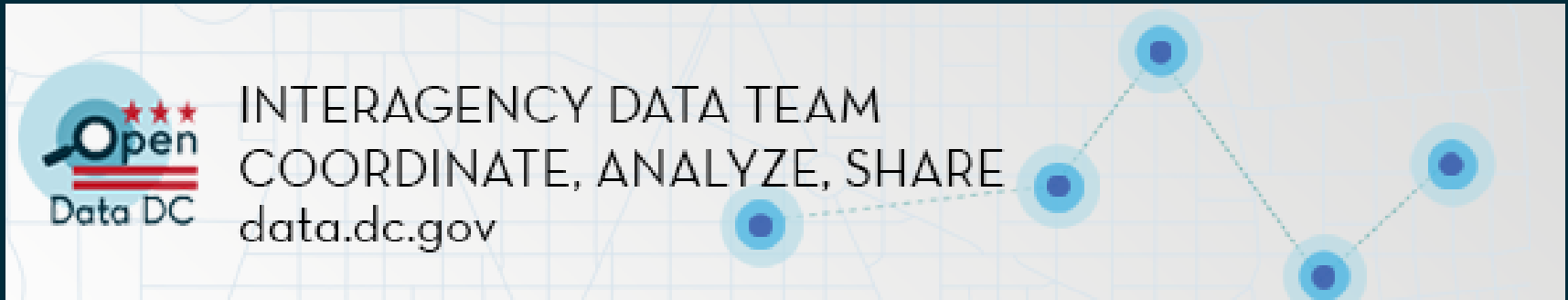




CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

Interagency Data Team

November 6, 2018

A horizontal banner with a light blue grid background. On the left is the 'Open Data DC' logo, which includes a magnifying glass icon, three red stars, and the text 'Open Data DC'. To the right of the logo, the text reads 'INTERAGENCY DATA TEAM', 'COORDINATE, ANALYZE, SHARE', and 'data.dc.gov'. On the right side of the banner, there is a network diagram consisting of five blue circular nodes connected by dashed lines.

Open Data DC
INTERAGENCY DATA TEAM
COORDINATE, ANALYZE, SHARE
data.dc.gov



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

WELCOME, AGENDA DATA TEAM NEWS

Michael Bentivegna
Data Visualization and Analysis Program Manager
Office of the Chief Technology Officer



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

AGENDA

- **Welcome, News & Updates**

Michael Bentivegna, Data Visualization and Analysis Program Manager
Office of the Chief Technology Officer

- **Data Science Working Group**

Peter Casey, Senior Data Scientist
Office of the Chief Technology Officer, Lab @ DC

- **FY19 Agency Data Submission Process for Level O Datasets**

Mario Field, Data Curation Program Manager
Office of the Chief Technology Officer

- **Enterprise Data Inventory – A Preview of New System in Testing**

Michael Bentivegna, Data Visualization and Analysis Program Manager
Office of the Chief Technology Officer



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BUSINESS INTELLIGENCE SOFTWARE UPDATES

- **Tableau Production version 2018.1.1**
 - Google sheets as a data source in production
- **Major upgrade to move to 2018.2**
 - Hold off on Desktop upgrades
- **Architecture review session**
- **Tableau Public w/automated ETL updates**
 - Hyper not currently supported by Informatica Tableau extension at this time
 - Box plug-in will be tested as a public data source



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BUSINESS INTELLIGENCE SOFTWARE UPDATES

- **MicroStrategy Production Version 10.11**
 - Moving MicroStrategy Web servers from Linux to Windows (WebLogic to IIS)
 - Creating public facing MicroStrategy web servers
 - Developing MicroStrategy integration guidelines for imbedding analytics in Agency websites and SharePoint.
 - New physical servers with more CPU power and memory is being deployed as the MicroStrategy Production servers



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

GIS SOFTWARE

DC AGOL - ARCGIS ONLINE/PORTAL

- **ArcGIS Server Upgraded to version 10.6**

- Some WMS map service format issues reported in Tableau

- **ArcGIS Portal version at 10.6.1**

- Reach users by sharing maps imbedded in web pages and through web applications
- Easily Add DC GIS Data
- **Leverage ready to use apps, Story Maps, Survey tools and data collection applications.**
- Use custom templates
- Easy access to Mobile
- Create Map/BI Galleries
- Data Initiatives Pages (<http://opendata.dc.gov/>)
- Hybrid of Cloud/DC GIS Infrastructure
- Included with citywide Esri ELA

<http://dcgis.maps.arcgis.com>

<https://dcgis.dc.gov/>





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BI TRAINING

- **Tableau Intro & Intermediate** (Eventbrite - search Tableau)
 - Introductory training
 - December 5 and 6 (FULL)
 - Intermediate training
 - December 12 and 13 (two seats left)
- **MicroStrategy Intro Classes** (Eventbrite - search MicroStrategy)
 - November 27 - 28 (FULL)
 - December 18 - 19 (eleven seats left)



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

GIS TRAINING

- **GIS** (PeopleSoft - request training - Search by name "GIS" or course number)
 - 230 Overview DC GIS Services February 6
 - 233 ArcGIS Online January 16



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

DATA SCIENCE WORKING GROUP

Peter Casey, Senior Data Scientist
Office of the Chief Technology Officer
Lab @ DC





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

DATA SCIENCE @ DC

- A monthly meeting of DC government employees interested in data science hosted by OCTO and The Lab @ DC
- **Topics:**
 - State of Data Science in DC Gov and as an industry
 - Issues in Data Science, such as ethics and fairness
 - Challenges to Data Science in DC Gov
 - Projects across DC Gov
- **Next Meeting:** Wednesday, November 14
- **Interested?** Peter.Casey@dc.gov



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

OPEN DATA THAT IS NOT ON THE OPEN DATA PORTAL...YET!!!

Mario Field, Data Curation PM
Office of the Chief Technology Officer





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

COUNT BY AGENCY (SAMPLE)

Agency	Document	Open Data	Num
DDOT		2	19
DMV			13
DPR			13
OVSJG			8
OCTO		13	8
DOH		1	7
DCPS			7
DCRA		1	7
DGS			5
DBH	3	1	4
OCP	1	2	4
DPW		4	4
DCOZ		2	4
DCPL			4

- **Numbers**

TOTAL - 189

DOCUMENT - 15

DUPLICATE - 32

ELIGIBLE - 142

- **Records marked as Documents and Duplicates are cursory OCTO review.**

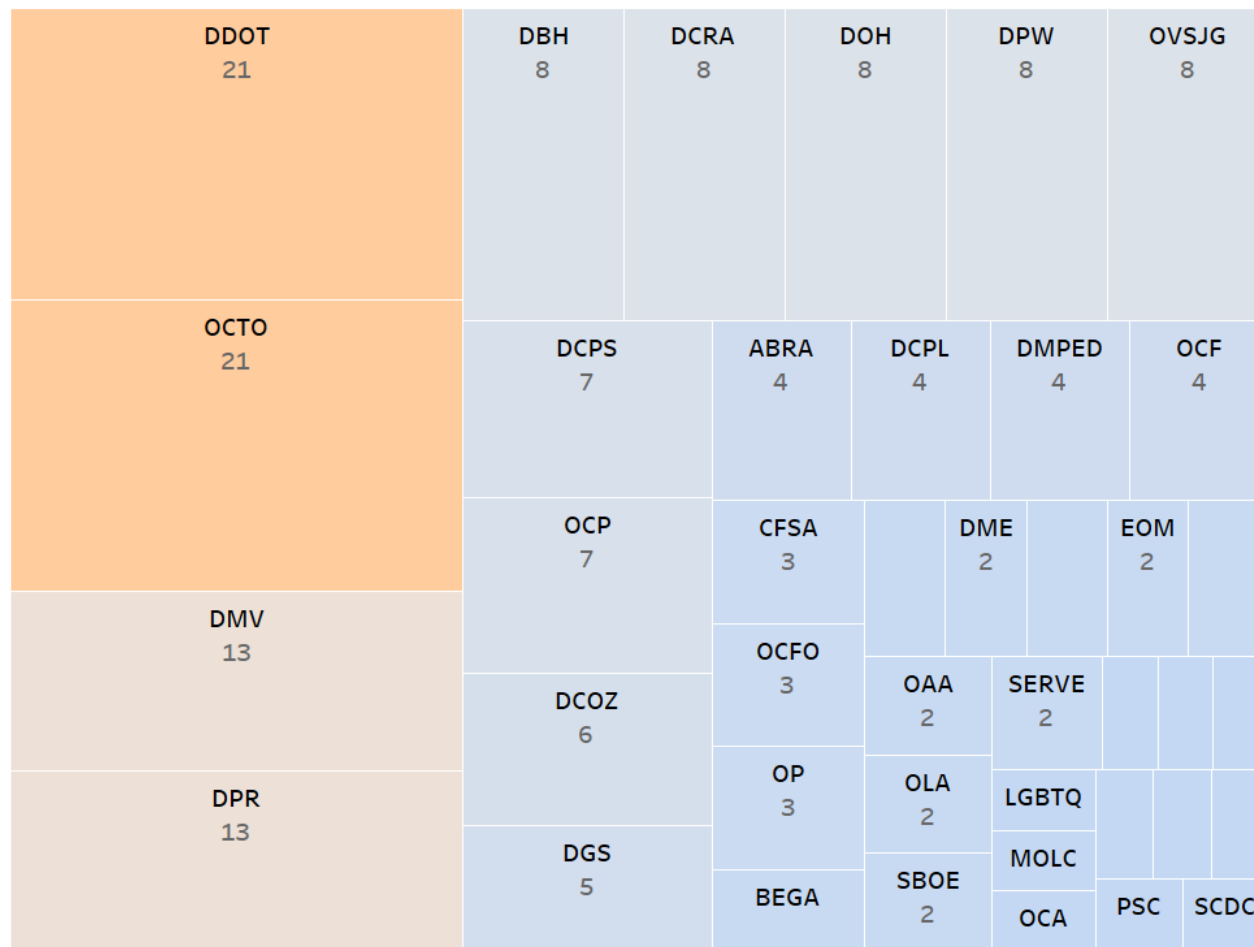
- **Covers 40 Agencies and Offices**



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BY AGENCY (TOTAL)

Agency Open Datasets





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BY CATEGORY THEME

Dataset Category	
Dataset Category	
Government Operations	38
Transportation	34
Health	19
Business and Economic D..	14
Public Safety	12
Education	12
Public Services	10
Planning Land Use and Zo..	9
Aerial Photography and S..	9
Location	7
Administrative and Other ..	7
Recreation	5
Utility and Communication	4
Property and Land	2
Financial	2
Facility and Structure	2
Environment	2
Historic	1

Data Submission Guide – Draft - Briefing

NOVEMBER 6, 2018

OCTO Data Division – Curation Program



BACKGROUND

1. Provides transparency and accountability for DC's data program
2. Allows agencies and the public to understand the overall data curation process.
3. Documents the requirements and guidelines needed for submission to ensure the integrity and consistency of data
4. Balances openness and transparency with privacy and security protections

Research

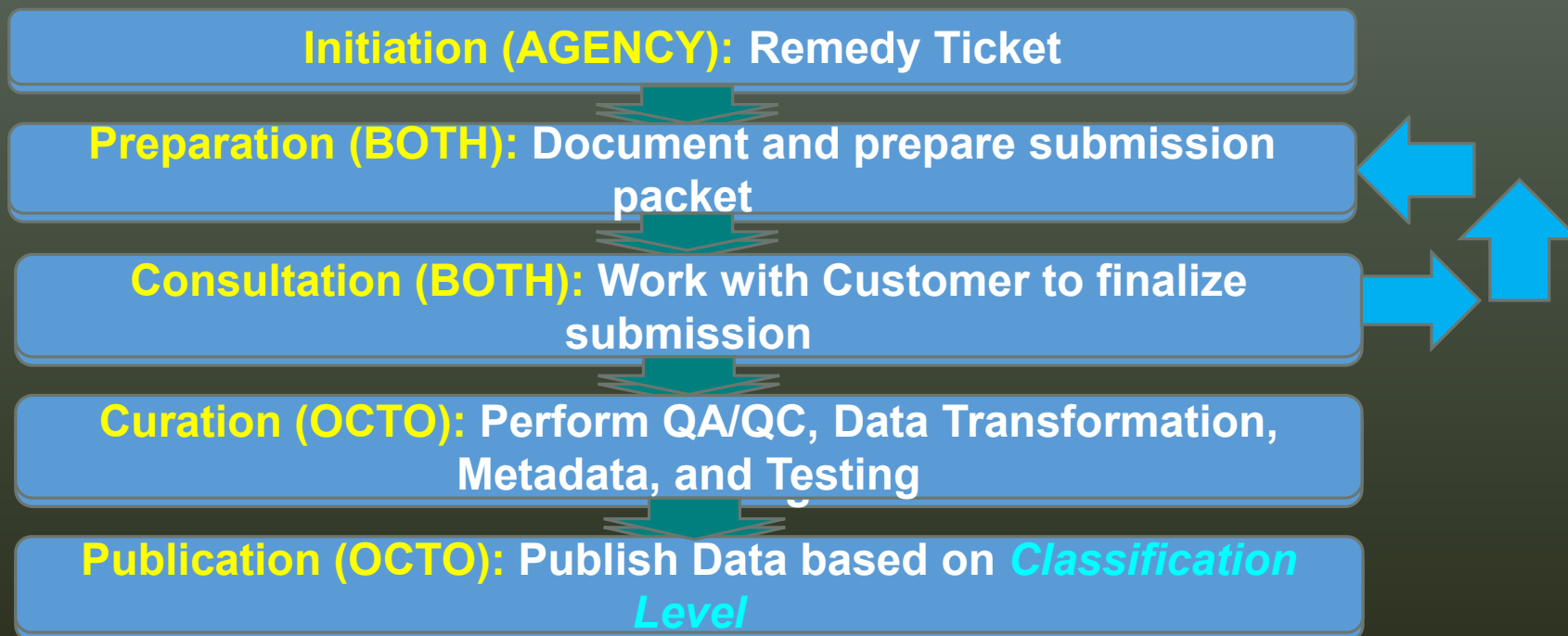
Review other jurisdictions

- New York State
- New York City
- Chicago
- Seattle
- San Francisco
- Maryland

Findings

- New York State has the most comprehensive documentation (roles, responsibilities, processes, dataset rules, etc.)
- Others provided more overall/general processes and steps.

OVERVIEW



INITIATION

Request For	Mario Field
Phone	(202) 727-1761
Email	mario.field@dc.gov
Requested By	Mario Field

GIS/BI Data and Visualization - Web Map and Agency Data Publishing

Requester Information

Is the phone number listed above correct?

Please select your Agency/Department/Location (use search to find location or address)

Request Details

Publish new, refresh or agency data

Please click the Submit button to create your service request

Add Attachments and Notes:

Ticket is closed once it is assigned to a data analyst

PREPARATION

Data Submission Packet (DSP)

- Remedy Ticket
- Data Submission Form
- Data Dictionary
- Data Mapping Document
- Approval Form
- Data Review Summary Report
- Additional Documentation

CONSULTATION

- Review DSP with Agency

CURATION

- Data Processing
- QA/QC
 - Data Rules
 - Data Format (5)
 - Data Field (12)
 - Addressing (3)
 - Miscellaneous (3)

- Metadata
 - Description
 - Topic and Keyword
 - Citation
 - Resource Details
 - Resource Maintenance
 - Resource Constraint
 - Spatial Reference
 - Resource Quality
 - Resource Lineage
 - Distribution
 - Entity Attribute
 - Metadata Reference

SECURITY

Database level – When data is loaded into a database, it gets placed in a secure schema with separate login credentials. Furthermore, the database can be encrypted to protect content.

Map service level – When the data is published in a map service, the software allows for security roles and users. The agency would define the role and provide a list of users and applications that can access the data.

Data Processing level – There are ways to process the data that create derivatives that mask individuality. Aggregation is the most common where data is presented in geographic areas or blocks that simplify and veil the data. Summarization is also a way to anonymize the data by combining and presenting individual records as groups or summaries.

Encryption - In select cases, encryption can be used in ETL development.

Other – OCTO possesses several other tools to customize and secure data. This includes Service-Oriented Architecture (SOA) and Application Programming Interface (API) Gateway

APPROVAL FORM

- Privacy and Security Concerns
- Proposed Mitigation
- District Agency
 - ADO and General Counsel
- OCTO
 - Data Curation Program Manager and General Counsel
- Changes Applied to address concerns

Appendix C – Data Submission Approval Form

DATASET TITLE

POTENTIAL PRIVACY OR SECURITY RISKS OR SENSITIVITY ISSUES

PROPOSED MITIGATION

District Agency or Office

Agency Data Officer: _____ Date: _____
Agency Director (level 2): _____ Date: _____
General Counsel: _____ Date: _____

Office of the Chief Technology Officer

Data Curation Program Manager: _____ Date: _____
Chief Technology Officer (Level 2): _____ Date: _____
General Counsel: _____ Date: _____

Recommendation:

Approved for publication
 Changes necessary

CHANGES NEEDED

CHANGES MADE

Department Agency Data Officer: _____ Date: _____

Data Curation Program Manager: _____ Date: _____

PUBLICATION ADDITIONAL STEPS

- Dataset Review
- Archiving
- Error Submission

ACTION ITEMS

- Develop a formalized, documented data curation workflow
- Finalize approval form and workflow
- Create an online data submission entry form in the same application that manages the EDI and integrate the two
- Develop training materials and a course describing this process
- Create an online mechanism for the public and agencies to submit correction requests



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

ENTERPRISE DATA INVENTORY PREVIEW NEW SYSTEM IN TESTING

Michael Bentivegna, Data Visualization and
Analysis Program Manager
Office of the Chief Technology Officer





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

ENTERPRISE DATA INVENTORY

Mayors Order 2017-115 : 2018-115

- Requires Agencies to annually review and update their Enterprise Data Inventory
- New application developed to review and update inventory
 - Application tested and in production: on or before Dec. 1th
- Agency inventory is due, reviewed by council ~~Dec. 11th~~ → Feb. 11th

The screenshot shows a web browser window with the URL opendata.dc.gov. The page features the DC.gov logo and a 'Sign In' link. The main content is titled 'Data Initiatives' with the subtitle 'Learn how DC initiatives are creating data to be more effective and efficient for everyone.' There are two featured cards:

- Cataloging Datasets**: A card with a background image of server racks. The title is 'Chief Data Officer's Report'. The text reads: 'A evaluation of the previous fiscal year's progress in implementing the Data Policy and publication of the Enterprise Dataset Inventory. 79 agencies recorded 1,640 enterprise datasets - nearly 80% of 99 agencies asked to participate. [Read full report.](#)' It includes a three-star rating and the text 'Publicly acknowledging District datasets.'
- Aerial Imagery & LiDAR**: A card with a background image of an aerial view of a city. The text reads: 'DC collects Orthophotography and now LiDAR data on a regular basis. These aerial images have a variety of uses including web based visualizations and to create derivative data such as buildings, roads and elevation models.' It includes a three-star rating and the text 'Preview and download the latest 2017 capture.'





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.



Enterprise Data Inventory



Michael Bentivegna

Select Agency : Office of the Chief Technology Officer

- Data Inventory
- Data Center
- Application
- Summary

Show 10 entries

Search:

	Dataset Name	Application Name	Dataset Classification	Dataset Classification Reason	Data O
	<lt;theheh->	Services	Level0	<lt;theheh->	agency
 	the dATASET	Google Analytics	Level 2: For District Government Use	KFHGXFGHG	the gov
	Telecommunications Request and Approval Tracking		level4	The request include critical infrastructure information about building network infrastructure changes, names of DC Government staff who are traveling and travel dates/locations, equipment orders, etc.	OCTO-
	SQA Testing Data for OCA Application Life Cycle Management		level2	This dataset is subject to one or more FOIA exemptions.	OCTO-





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.



THE DISTRICT OF COLUMBIA

Enterprise Data Inventory

Select Agency :

Office of the Chief Technology Officer

Data Inventory

Data Center

Application

Show 10 entries

	Dataset Name	Application Name	Dataset Classification
	<lt;theheh->	Services	Level0
	the dATASET	Google Analytics	Level 2: For District Government Use
	Telecommunications Request and Approval Tracking		level4
	SQA Testing Data for OCA Application Life Cycle Management		level2
	Avaya Experience Portal Call Statistics		level1

CLOSE

Notes

Add a new note

Update

Updated By:

Michael Bentivegna

Updated Date:

11/6/2018 12:00:16 PM

Comments :

Revised by OCTO Liaison

Updated By:

Priya Kaja

Updated Date:

11/6/2018 11:52:13 AM

Comments :

Dataset is updated

Updated By:

Priya Kaja

Updated Date:

11/6/2018 11:40:15 AM

Comments :

Dataset is updated

Updated By:

Priya Kaja

Updated Date:

11/6/2018 11:38:45 AM





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

OPEN DISCUSSION

NEXT MEETING SCHEDULED FOR
NEXT YEAR

THANK YOU 😊

