



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.



INTERAGENCY DATA TEAM
COORDINATE, ANALYZE, SHARE
data.dc.gov

May 17, 2017

Hosted by the Office of the Chief Technology Officer

octo.data@dc.gov





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

AGENDA

- Welcome, Data Program News & Updates
Barney Krucoff, Chief Data Officer, Office of the Chief Technology Officer
- Sunlight Foundation Remarks on Open Data
Alex Howard, Deputy Director
- Data Policy Implementation, How to Measure Success
Barney Krucoff
- The Lab @ DC's Security Procedures & Templated Data Sharing Agreements
Kevin Wilson, Senior Data Scientist, Office of the City Administrator
- Open Discussion on Data Sharing Agreements
Barney Krucoff

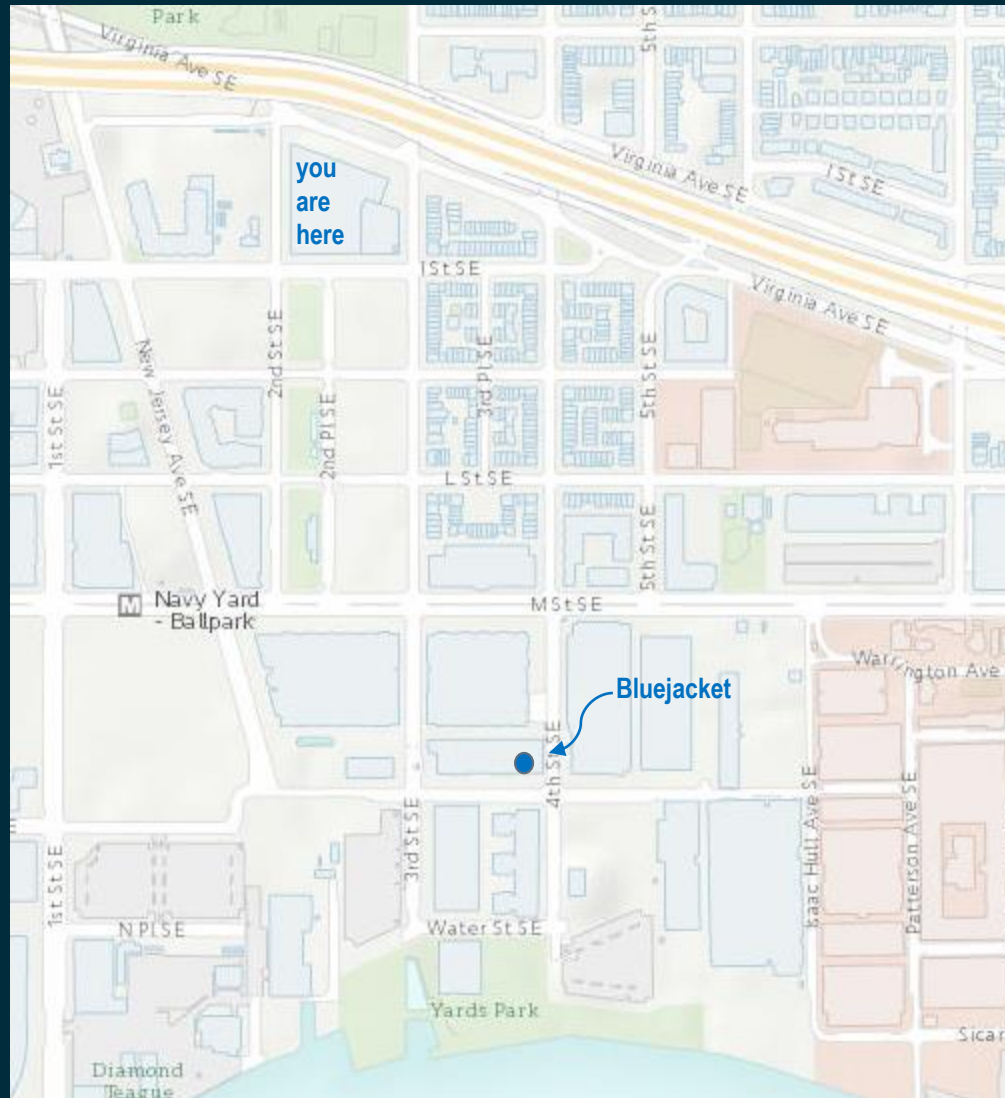
Unofficial Celebration of New Data Policy
Bluejacket, 300 Tingey Street SE





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

Unofficial Celebration of New Data Policy
Bluejacket, 300 Tingey Street SE





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

OPEN DATA POSTINGS

- Best Management Practice (DDOE)
- Bicycle Lane
- Car Share Location
- Clean Team Service Area
- DC Land
- Early Polling Place
- Liquor License
- Main Street
- Notary Public
- Owner Point (shows parcels not in VPM)
- Park and Recreation Area
- Polling Place
- Residential Parking Permit Block
- Summer Crime Initiative Area
- Triangle and Median
- 311 Service Request, Crime, DDOT Permit, BBL, and Building Permit have been converted to point layers for performance.



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

NEXT OPEN DATA POSTINGS

- Cooling Center
- Library
- Parking Meter
- Recreation Center
- Senior Service Network Participants
- Ward



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

REMARKS ON OPEN DATA AND DC'S NEW DATA POLICY



Alex Howard
Deputy Director
Sunlight Foundation





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

DATA POLICY IMPLEMENTATION HOW TO MEASURE SUCCESS

Barney Krucoff





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

WHAT'S NEW IN POLICY

- Copyrighted material exemption
- Legal review of every enterprise dataset classification
- Signature 4/27/2017!
- Data policy can be found at:
<https://octo.dc.gov/page/district-columbia-data-policy>



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.



Open Data



Protected Data

two sides of the same coin



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

BALANCED GOALS

- The data created and managed by the District government are **valuable assets** and are **independent of the information systems** in which the data reside....
- The greatest **value from the District's investment in data can only be realized when** the enterprise datasets are **freely shared** among District agencies, with federal and regional governments, and with the public to the fullest extent consistent with safety, privacy, and security....
- Inappropriate disclosure of personal information and misuse of data for activities such as identity theft are significant concerns. To **protect the safety, privacy, and security** of residents, workforce members, clients, partners, stakeholders, visitors, and others, datasets requiring protection shall be identified and....



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

WITH REGARD TO ENTERPRISE DATASETS, THE DISTRICT SHALL

- Maintain an inventory of its enterprise datasets;
- Classify enterprise datasets by level of sensitivity;
- Regularly publish the inventory, including the classifications, as an open dataset;
- Openly share some enterprise datasets while carefully protecting others.



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

IMPORTANT DEFINITIONS

- “Data” is a very expansive definition.
- “Dataset” is more limited.
- “Enterprise Dataset” is what requirements relate to and is much more restrictive.



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

ENTERPRISE DATASET INVENTORY TIMELINE

HOW TO GET STARTED

- **May 4:** Mayoral Order 2017-115, District of Columbia Data Policy Announced
- **June 5:** AISO and ADO selection process complete and reported to OCTO. (Section IV B and D)
- **Week of June 5:** OCTO Data Team liaisons will start actively contacting ADOs* (**need one volunteer**)

- **November 28:** Agency contributions to the Data Inventory must be complete.
- **January 26, 2018:** Data Inventory is published as Level O, Open
- **Future,** updated data inventories will be published on November 1 of each year.



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

YOUR AGENCY MAY BE HALFWAY THROUGH YOUR INVENTORY

What the DC Council typically asks:

- Please identify all databases maintained by your office, including the following:
 - A detailed description of the information tracking within each system;
 - The age of the system and any substantial upgrades that have been made or planned to be made;
 - How the information is protected; and
 - Whether the public can be granted access to all or part of each system.



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

ANNUAL INVENTORY & CLASSIFICATION

Dataset classification levels are defined as the following:

- Level 0, Open,
- Level 1, Public, Not Proactively Released
- Level 2, For District Government Use
- Level 3, Confidential
- Level 4, Restricted Confidential



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

OTHER SECTIONS

- PERSONNEL AND ROLES
- ENTERPRISE DATASET INVENTORY, CLASSIFICATION, AND PRIORITIZATION
- MINIMUM DATA PROTECTION STANDARDS
- DATA CATALOGS (opendata.dc.gov)
- STREAMLINED PROCESSES FOR INTERAGENCY DATA SHARING
- NEXUS BETWEEN FOIA AND LEVEL O, OPEN DATASETS
- LEVEL O, OPEN, DATA LEGAL POLICY, AND LICENSING



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HOW WILL WE MEASURE SUCCESS?

Criteria	Question/Answer
Executive Sponsorship	
Has the Mayor signed the policy?	Yes! 4/27/2017
Has the Chief Technology Officer named a CDO?	Yes! Barney.Krucoff@dc.gov
Has the Chief Technology Officer named a CISO?	Yes! John.MacMichael@dc.gov
Is OCA reinforcing policy through the Stat and KPI processes?	Yes/No
Public Body Participation	
Have Public Body Directors designated ADOs?	# of #? Mayoral bodies # of #? Independent bodies
Have Public Body Directors designated AISOs?	# of #? Mayoral bodies # of #? Independent bodies



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HOW WILL WE MEASURE SUCCESS?

Criteria	Question/Answer
Enterprise Dataset Inventory	
Did OCTO release the data inventory tool within 60 days?	Yes/No
Has OCTO released a list of eligible and mandated public bodies?	Yes/No
Has OCTO provided public bodies training on the inventory process?	# of # eligible bodies
Has OCTO publicly released the Enterprise Dataset Inventory within 270 days of the release of the inventory tool?	Yes/No
Is that inventory being maintained during the year between major releases?	Yes/No
Inventory breadth: How many public bodies have listed at least one dataset?	# of # eligible bodies # of # mandated bodies
What public bodies have not inventoried at least one dataset?	Listing
Inventory depth: What is the number of datasets inventoried?	#
Percent of all Inventoried Enterprise Datasets reviewed by General Counsels or their designees.	%
Are there observed gaps in the data inventory?	Listing of gaps in inventory produced by OGAG



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HOW WILL WE MEASURE SUCCESS?

Criteria	Question / Answer
Enterprise Dataset Classification	
What is the breakdown of datasets by classification 0 to 4?	Pie Chart # and %
Is the percentage of Level 0 datasets increasing year over year?	Level 0 / Total compared to previous year as of Oct 1.
Complementary Relationship Between Open Data and FOIA	
Percent of eligible agencies with accounts on tracking/processing system (currently FOIA Express)	# of #
Are there datasets that have been holistically released under FOIA not classified Level 1 or below?	Listing
Is FOIA request metadata being released as open data in a timely manner as prescribed by the policy?	Yes/No
Is FOIA request metadata being released as open data in a safe manner respectful of privacy and security concerns?	OGAG Opinion
What datasets have been classified as level 1 and stated reason?	Listing



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

THE LAB @ DC'S SECURITY PROCEDURES & TEMPLATED DATA SHARING AGREEMENTS

Kevin Wilson





THE **LAB** @ DC



The Lab @ DC's Template Data Use Agreement and Security Policy

Kevin Wilson, Senior Data Scientist

Section I

INTRODUCTION

What is The Lab @ DC?

- New group in Office of the City Administrator
- Bring academic findings into government
- Help agencies measure the effects of policy changes

Example Projects of The Lab

RCTs



Nurse Triage Line



Body Worn Cameras

Rapid RCTs



457(b) Enrollment



TANF Recertification

Predictive
Modeling



Risk-Based Proactive
Housing Inspections



LEA Payments

Common thread....

- All of our projects require data...
- ...often from multiple agencies...
- ...often of a sensitive nature.

New Data Policy: What it does

- Establishes data taxonomy based on sensitivity and some legal requirements
- Describes general transparency and sharing principles
- Describes general security principles

New Data Policy: What it doesn't do

- Template interagency data use agreements
- Explicit security protocols for data
- Publication and disclosure protocols

Why not?

- Both security and DUAs require tailoring to project needs

Examples are Nice Though

- The Lab @ DC has built a template Data Use Agreement and Security Policy
- Rest of talk is about principles underlying these documents

Section II

THE LAB @ DC'S SECURITY POLICY

Data Levels

- Level 0: Open Data
- Level 1: Public data upon request
- Level 2: Proprietary
- Level 3: Confidential
- Level 4: Restricted

Main Points for Sharing and Security

- Levels 0 and 1
 - No extra security required
 - Proactive sharing with other agencies encouraged
- Levels 3 and 4
 - Extra security required
 - Sharing should be governed by DUA
- Level 2
 - Depends on nature of data

The Lab's Policy Principles

- Treat all data as Level 3 data
- Level 4 data requires special negotiation
- Look to HIPAA (probably strictest regulations) for guidance

The Lab's Policy Principles

- Access Controls
- Encryption
- Deletion
- Employee Restrictions

The Lab's Policy Principles

- Access Controls
 - Two-factor authentication whenever possible
 - (Mostly) centralized computation
 - Internal data sharing requires explicit authorization

The Lab's Policy Principles

- Encryption
 - At rest
 - Full disk encryption for all Lab computers
 - Full disk encryption of data disks in Cloud
 - In transit
 - All transactions performed over SSL/TLS
 - Strict nginx config limits algos

The Lab's Policy Principles

- Deletion
 - After a project is completed, The Lab deletes Level 3 data.
 - Exceptions:
 - Replicability requirements
 - Anonymization

The Lab's Policy Principles

- Employee Restrictions
 - All access to resources over VPN
 - Avoid accessing Level 3 data on open wifi
 - WPA2 encryption on home network

Software Supporting Principles

- JupyterHub
- Secure Uploader
- GitHub (handles 2FA)

Section IV

TEMPLATE DATA USE AGREEMENT

Main Features

- Explicit enumeration of purposes of sharing
- Data security policy
- Data custodian
- Code to produce data shared along with data
- Unauthorized disclosure provisions
- Publication Restrictions

Main Features

- Unauthorized disclosure provisions
 - Defines OCA's responsibilities when data is disclosed not pursuant to agreement
 - Requires OCA to make reasonable efforts to mitigate effects

Main Features

- Publication Restrictions
 - No raw Level 1+ data
 - Agency reviews analysis plan
 - Open Science Framework

In Flight With...



THE **LAB** @ DC



DC | DEPARTMENT of
HUMAN SERVICES



DISTRICT OF COLUMBIA
DFS
DEPARTMENT OF
FORENSIC SCIENCES

GOVERNMENT OF THE DISTRICT OF COLUMBIA
DOH
DEPARTMENT OF HEALTH
Promote. Prevent. Protect.



THE **LAB** @ DC



THANK YOU!

Kevin Wilson

kevin.wilson@dc.gov

The Lab @ DC

Office of the City Administrator

Executive Office of the Mayor



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

OPEN DISCUSSION DATA SHARING AGREEMENTS

Barney Krucoff





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

THANKS FOR JOINING US 😊

Feedback

open.data@dc.gov

APPENDIX

Data Levels

- Level 0: Open Data
 - Should be proactively put onto opendata.dc.gov
 - Examples:
 - Certificates of Occupancy (DCRA)
 - Basic Business Licenses (DCRA)
 - 311 Calls (OUC)
 - Integrated Tax System Public Extract (OTR)

Data Levels

- Level 1: Public data upon request
 - Public, but not proactively disclosed
 - Examples:
 - Voter files (DCBOEE)

Data Levels

- Level 2: Proprietary
 - Not subject to FOIA and not sensitive
 - Examples:
 - Data sets purchased by agencies (e.g., Morningstar)

Data Levels

- Level 3: Confidential
 - Not subject to FOIA and sensitive
 - Examples:
 - Students' records (FERPA/COPPA)
 - Legal determinations (Attorney-client privilege)
 - Payment card information
 - (Usually) Personally Identifiable Information

Data Levels

- Level 4: Restricted
 - Could potentially lead to major damage or injury
 - Examples:
 - VIN numbers of unmarked police vehicles
 - Names of witnesses in ongoing criminal trials

Main Features

- Explicit enumeration of purposes of sharing
 - Purposes of data determined in advance
 - Any further use requires negotiation

Main Features

- Data security policy
 - Binds all Lab employees to the security policy

Main Features

- Data custodian
 - Specific member of The Lab
 - Responsible for data security and integrity of the data set

Main Features

- Code to produce data shared
 - Without code, difficult to understand provenance of data