

CHIEF DATA OFFICER ANNUAL REPORT 2020

WEARE GOVERNMENT OF THE DISTRICT OF COLUMBIA CMURIEL BOWSER, MAYOR

Table of Contents

Introduction	6
Where We Are Today	8
Enterprise Data Inventory Report	8
FOIA Tracking	15
2019 Accomplishments	18
Where We Are Going: Draft Tech Plan and Progress on 2019 Goals	32
Conclusion	37

Contributors

Enterprise Dataset Inventory, Data Curation and Analysis

Michael Bentivegna Program Manager, Data Visualization and Analysis

Eva Reid Senior Analyst and Enterprise Data Inventory Lead

Mario Field Program Manager, Data Curation

Tim Harwood Senior BI Analyst **Open Data Portal Initiatives Page**

Alexandre Santos GIS IT Specialist

Editing and Design

Nina Liggett Deputy Communications Director



Author Acknowledgments

This is the District of Columbia's Chief Data Officer's Annual Report produced within the of context the draft strategic plan for the Office of the Chief Technology Officer (OCTO).

The Draft Technology Strategic Plan for DC: Unleashing the Possible, though still a draft, is providing direction and clarity to the agency, the District government, and the Chief Data Officer (CDO) program. OCTO is building on a strong foundation and expanding services to help DC Government by providing agencies opportunities to share, store and analyze data to uncover insights to improve services for residents; and the development of this plan has brought renewed energy and mission to OCTO, thus "unleashing the possible."

I would like to acknowledge the Agency Data Officers and Agency General Counsels for their hard work gathering and preparing the Enterprise Dataset Inventory (EDI) for publication as well as the Open Government Advisory Group.

Finally, a special thanks goes to the staff and contractors of the OCTO Data Team.

Bay Kin

Barney Krucoff Chief Data Officer District of Columbia Office of the Chief Technology Officer

In April 2017, Mayor Muriel Bowser issued Mayor's Order 2017-115, the "District of Columbia Data Policy" ¹ (Data Policy), with the goal of leading the District of Columbia government toward more open and efficient use and sharing of government data.

The policy established these principles acknowledging the value of data to the District and the inherent need to balance openness:

- Data are valuable assets independent of the information systems in which the data reside.
- The greatest value from those assets is realized when they are freely shared to the extent consistent with the protection of safety, privacy, and security.

In June 2018, Mayor Bowser issued Mayor's Order 2018-050, "District of Columbia Data Policy,"² aligning the publication of the District's Enterprise Dataset Inventory (EDI) and this annual report with the national observance of Sunshine Week. Sunshine Week is a national initiative spearheaded by the American Society of News Editors to educate the public about the importance of open government.

The purpose of the CDO annual report is to report accomplishments and EDI results; assess current progress; as well as share where we are going and set new goals.

Introduction

Released November 2019, the <u>Draft Technology</u> <u>Strategic Plan for DC: Unleashing the Possible</u> (Draft Tech Plan) consists of four key components:

- **1. Vision:** Unleashing the possible for DC in the digital age.
- 2. Mission: Empower DC government through technology by providing valued services, advising agencies, and collaboratively governing information technology.
- **3. Commitments:** Nine commitments to our agencies, residents, businesses, visitors, and to the OCTO team. They focus on providing exceptional services, products, technology guidance and governance.
 - Relevant to this report is Commitment 4: Manage, organize, and leverage data to facilitate DC government mission delivery.
- **4. Initiatives:** The tasks and initiatives OCTO will undertake to ensure DC agencies have the tools to manage, organize, and leverage data to make better decisions.

Commitment 4 Initiatives

- **4.1** Identify high-value mission use cases to improve the use of data in service of specific Mayoral initiatives and focus areas.
- **4.2** Streamline policy and process to share and use data.
- **4.3** Support agency use of big data and artificial intelligence

As Chief Data Officer it is my job to help the District of Columbia government realize the greatest value from the data it collects and manages. The Tech Plan empowers my team and DC Government to do just that by driving us to provide new services and to expand our technology offerings.

Initiative 4.1inIdentify high-valueAccmission use cases toaccimprove the use of datainin service of specificMayoral initiatives and

Initiative 4.2

focus areas.

Streamline policy and process to share and use data.

Initiative 4.3

Support agency use of big data and artificial intelligence Section 2 of this report expands on achievements in data in the priorities areas identified in Mayor Bowser's 2020 Accountability Report ³:

- 1. Good Government
- 2. Housing
- 3. Education
- 4. Jobs

- 5. Transportation
- 6. Public Safety
- 7. Health
- 8. Environment

This initiavtive closely aligns with Mayor's Order 2017-115, the District of Columbia Data Policy, for example, the annual Enterprise Dataset Inventory, a mainstay of every annual CDO report, is all about knowing and classifying what data the District holds, which is a prerequisite for leveraging and sharing that data.

Directing us toward emerging technologies, we have been working on big data for a couple of years and will provide an update on the DC Data Lake in this report. Artificial intelligence along with other Smart City technology will be a largely new venture for the OCTO Data Team, as we expand our commitments to being the data champions of the District.

^{1. &}lt;u>https://www.dcregs.dc.gov/Common/NoticeDetail.aspx?noticeId=N0063620</u>

^{2. &}lt;u>https://www.dcregs.dc.gov/Common/NoticeDetail.aspx?noticeId=N0071235</u>

^{3. &}lt;a href="https://mayor.dc.gov/page/mayor-muriel-bowsers-2020-accountability-report/">https://mayor.dc.gov/page/mayor-muriel-bowsers-2020-accountability-report/

Where We Are Today: The 2020 Enterprise Data Inventory Report, FOIA tracking and 2019 Accomplishments

Enterprise Dataset Inventory (EDI)

The Data Policy mandates that public bodies in the District government create and maintain an EDI under the leadership of OCTO. The EDI requires agencies under the direct authority of the Mayor to record any "enterprise dataset," which is "a dataset that directly supports the mission of one or more public bodies." The Data Policy also requests that independent District government agencies, not under the Mayor's authority, participate in the EDI, though their participation is not required. What follows is an analysis of the 2020 EDI. The inventory is available through the city's Open Data Portal.⁴



Figure 1

How many agencies participated?

82 agencies recorded 1,915 enterprise datasets in 2020, up from 1,779 datasets recorded by 75 agencies in 2019.

Several smaller agencies are consolidated under the Executive Office of the Major (EOM) for purposes of the inventory. The number of independent agencies increased to 16 from 11 the previous year.⁵



Several agencies increased the number of enterprise datasets they recorded. We are particularly proud of the Office of the Deputy Mayor for Education (DME), which added 35 records through the creation of their EdScape application (edscape.dc.gov). Other agencies notably contributing to the growth of the inventory this year were the District Department of Transportation (DDOT), which recorded 34 additional records; the Office of Campaign Finance (OCF), 28 additional records; and the Office of Unified Communications (OUC), 10 additional records. Small agencies also made significant contributions, for example, the Department of Small Local Business Development (DSLBD), which added 9 datasets.

200

50

ОСТО DDOT DCPS DOH DOEE OP

> DME OUC

> > OCF

DPW MPD DHS DCRA OSSE ORM DPR DOES DHCF 0CP DMV DISB DCPL OAG DOC **CFSA** FEMS DBH DMPED DC07 **HSEMA** EOM DCWATER DCHR BOE DGS

> OCTFME DCOA

OVSJG DYRS

> CAH OHR

OCME OCA DFHV UDC PSC

SBOE RPTAC

OVA DDR

MOCA 🗖

DHCD 🖿 DFS

SERVE 📕

SCDC OSA

OLA 📕 OAH 📕

OAA 📘

MOTA DMHHS

BEGA

ABRA OTA 📘

OS 🛛 OPGS OLRCB MOLC LGBTQ DCPCSB DCHA CFMB

0

Number of Enterprise Datasets

150

100

9

250



Figure 3 How did agencies

Dataset classification system

Level 0 is Open. Level 0 data is any data that is open to the public and should be proactively released. This is the default classification for the EDI and applies to any dataset that agencies do not determine to have a higher security level. In 2020, 859 enterprise datasets (44%) were classified as Open, up from 794 in 2019. Open is the most common classification in the EDI.

Level 1 is Public but Not Proactively Released. Level 1 data is not protected from public disclosure but is not proactively published because of concerns over safety, privacy, security, or legal concerns. The number of datasets classified Level 1 increased in 2020 to 179 (9%) from 144 the previous year.

Level 2 is For District Government Use. Level 2 data is "subject to one or more FOIA exemptions, [but] is not highly sensitive and may be distributed within the District government." The number of datasets classified Level 2 increased to 219 (11%) in 2020 from 188 the previous year.

Level 3 is Confidential. This includes data that is "protected from disclosure by law" and that is either highly sensitive or legally restricted from disclosure to other public bodies. In 2020, 562 datasets (29%) were classified Confidential, down slightly from 566 the previous year. Confidential is the second most common classification in the EDI.

Level 4, Restricted Confidential. It is the rarest classification in the EDI. This refers to datasets for which "unauthorized disclosure could potentially cause major damage or injury, including death, or otherwise significantly impair the ability of the agency to perform its statutory functions." The number of datasets classified Restricted Confidential increased to 97 (5%) in 2020 from 87 the previous year.



Figure 4 Count of Sensitivity Flags

Figure 4 highlights the underlying reasons that datasets are classified as Level 3 or Level 4. In response to a recommendation from the DC Auditor, for the 2020 EDI Year, agencies were required to include an explanation of what type of Personally Identifiable Information (PII) or sensitive information these records hold as structured data. Previously this information had only been collected in narrative format.

Agencies are required to select all that apply to their datasets from a list of reasons/laws that restrict the sharing of data. Any one dataset may have many reasons it is Level 3 or Level 4.





The EDI identified 245 Level 0 datasets that are not in the Open Data Portal, an increase of 29 from 216 the previous year. OCTO will continue to work with agencies, the Open Government Advisory Group, and the community to prioritize posting these Level 0 datasets identified on the Open Data Portal.⁷



opendata.dc.gov

Office of the Chief Technology Officer | Governement of the District of Columbia

Agencies are also asked to categorize their datasets according to their contents and purpose. The number of enterprise datasets increased in several categories, especially Education (+37), Public Services (+34), Transportation (+30), Public Safety (+19), and Business and Economic Development (+11). The number of datasets only decreased slightly in a few categories including Utility and Communication (-5) and Government Operations (-3).

Figure 6

How the Agencies Categoriezed Their Data



Number of Enterprise Datasets

Retired Datasets

Datasets are retired if they are no longer in use for agencies' operations, duplicated, or otherwise improperly recorded. In 2020, 16 enterprise datasets were retired from the previous universe.⁸

Notably OCTO retired 13 datasets. Reasons for retirement are collected as part the inventory. In OCTO's case, typically another agency was the actual owner of a dataset that it had included in its inventory in previous years. No other agency retired more than 4 datasets. Retired datasets were widely dispersed by classification.

Figure 7

Which Agencies Retired Datasets



Figure 8

How Were Retired Datasets Classified?



Establishing and Analyzing a FOIA-Open Data Feedback Loop

The Data Policy strives to balance openness and security while prioritizing transparency. The policy states that "enterprise datasets shall be open by default, meaning that their existence will be publicly acknowledged, and further, if enterprise datasets are not shared, an explanation for restricting access will be publicly provided." In other words, "open by default" means that the District will publicly acknowledge all enterprise datasets, which is accomplished by publishing the annual EDI, although some may be labeled as confidential. The Data Policy envisions a "nexus between FOIA and level zero open datasets, where FOIA and Open Data are distinct but complementary practices," stating:

- FOIA request-tracking data should inform public bodies about public demand for open data. In fact, the Data Policy mandates use of the system for mayoral agencies.
- Open data publication of FOIA requesttracking data can help residents hold public bodies accountable for the timely and consistent processing of requests.
- Successful appeals for datasets previously denied under FOIA exemptions can inform public bodies about potential errors in dataset classification.

FOIA Part 1: Tracking and Reporting

As required by law, the Executive Office of the Mayor (EOM) publishes a District Government-Wide Cumulative FOIA Report annually.⁹ This year, for the first time, the detailed portion of the Annual FOIA Report was published as Open Data.¹⁰ Having the aggregated FOIA data in machine-readable format, as prescribed by the Data Policy, has been very helpful for analysis.

In addition to EOM aggregating the FOIA Data annually, the District has invested in FOIAXpress, an enterprise system which helps track and manage FOIA requests. The FOIAXpress system aids FOIA Officers in tracking, processing, and reporting on their work. It has been, and remains, a goal to increase the utilization of the FOIAXpress system within the government. Table 1 shows the growing utilization of the system.

Measures	2017	2018	2019
# of agencies enrolled in FOIAXpress	53	55	64
# of agencies processing at least one request in system	51	52	57
# of active DC government users (FOIA Officers)	112	141	153

Table 1: FOIAXpress Utilization by Fiscal Year

OCTO is focused on increasing the number of, and better training for, FOIAxpress users. In 2018 when comparing the EOM Cumulative Report to what was processed in FOIAXpress, I found that the number of FOIA requests is increasing even faster than use of the FOIAXpress system. I'm pleased to report that that trend reversed in 2019. The number of FOIA requests grew at a more modest pace, and use of the FOIAXpress system grew faster, gaining market share.

Table 2: Share of all FOIA Requests Processed Inside the FOIAXpress System

FOIA Requests	FY 2017	FY 2018	FY 2019	FY 2018-2019 Percent Change
Total reported by EOM	8,274	10,450	10,836	3%
Tracked in FOIAXpress System	6,762	7,943	8,701	9%
% of requests in FOIAXpress	82%	76%	80%	

It is also a goal to increase the percentage of requests tracked in FOIAXpress and increase the quality of that tracking. The 20% of requests occurring outside the system are concentrated in two categories: independent agencies that operate outside FOIAXpress entirely and Mayoral agencies that have the system but don't use it to track and process the majority of their requests.

Independent agencies with the most requests outside the system:

- DC Board of Elections
- DC Council
- DC Housing Authority
- DC Water
- Office of Police Complaints
- Office of the Chief Financial Officer
- Public Charter School Board

Mayoral agencies which do not have FOIAXpress. OCTO will reach out, provide licenses, and training:

- Department of Youth Rehabilitation Services
- Office of Disability Rights
- Office of Victims Services and Justice Grants

Agencies/offices that have and use FOIAXpress but aren't broken out in Mayoral agencies that do use FOIAXpress, but process most of their requests outside the system include the following:

- Alcoholic Beverage Regulation Administration
- District of Columbia Retirement Board
- Department of For Hire Vehicles
- Department of Human Services
- Department of Corrections
- Department of Small Local Business Development
- Office of Human Rights
- University of the District of Columbia

Agencies/offices rolled up into the reports of their parent organizations:

- Deputy Mayor for Public Safety and Justice (Office of the City Administrator)
- Office of Open Government (Board of Ethics and Government Accountability)

FOIA Part 2: Using FOIA to Find Data That Should Be Open

As CDO, I read through a sample of FY 2018 and FY 2019 FOIA requests (those that were in FOIAXpress) where "structured data" was specifically requested and the data granted in whole or in part. I also reviewed the annual FOIA Appeals Report. I was looking for instances where a requester was granted access to structured data not otherwise available on the Open Data Portal (opendata.dc.gov.) Of the datasets described in last year's report, we didn't make much progress.

The notable exception is data on dockless bikes and scooters. DDOT has made it a requirement that permit holders provide this data to the District. OCTO is helping DDOT capture, store, and analyze the data (see DC Data Lake later in this report). That said, there are privacy concerns, and we currently have the dockless trip data classified as Level 3.



Accomplishments from our Data Team and Agency Partners

*** * *** Good Government

What Works Cities Gold

Technical Capability: Open Data



Mayor Bowser and her team display the award for What Works Cities Gold certification on April 25, 2019.

What Works Cities, a program of Bloomberg Philanthropies, "evaluates how well cities are managed by measuring the extent to which city leaders incorporate data and evidence in their decision-making."¹¹ DC was one of seven cities nationwide to achieve Gold status. In addition to DC, the Gold-certified cities are: Arlington, TX; Kansas City, MO; Los Angeles, CA; Louisville, KY; Philadelphia, PA; and Scottsdale, AZ. No cities have yet achieved the highest level of certification, Platinum. The best thing is feedback that the certification process provides.

Figure 9



What Works Cities Criteria on Which We Scored Well and Where We Have Room to Improve

Helping District Systems Talk to Each Other

Technical Capabilities: API Gateway and Service-Oriented Architecture

To get the full value from the District's data, we must be able to move it around from application to application across agencies. There are safe and efficient ways to move data around the government and in and out of the cloud, and less safe and less efficient ways.

The use of Application Programming Interfaces (API) are becoming the universal way for transferring data to and from clients and backend systems. With an ever-growing number of APIs being put into use, the question of API management, security and access becomes more relevant. OCTO has invested in an API Gateway that can provide a solution to these questions. Serving as a multifaceted component for the OCTO Middleware team, the API Gateway promotes and supports microservices architecture with load balancing, lightweight translation, and access control. The API Gateway is supporting thousands of API calls for a number of District applications, including OUC 311 Cards, Metropolitan Police Department (MPD) Crime Cards and ShotSpotter, DC Public Schools MySchools, Office of Risk Management Enterprise Risk Management System (ERisk), and Child and Family Services Agency Community Portal.

As CDO, I strongly encourage agencies to employ OCTO's Middleware team and support technology to ensure their transactions move between systems securely and reliably. OCTO's Middleware Team completed several new integrations this year in addition to maintaining existing integrations with some of the District's largest systems, including the financial, procurement, and human resources systems.

Utilizing the Service Oriented Architecture (SOA) platform, the OCTO Middleware team assisted

MPD with the migration of their ShotSpotter platform into a cloud-hosted environment. ShotSpotter is a gunshot detection, acoustic surveillance technology that uses sensors to detect, locate, and alert law enforcement agencies of potential gunfire incidents in real time. Data collected through this platform is securely transferred through the SOA platform and is made available to additional MPD platforms for analysis and situational awareness.

The OCTO Middleware team continues to implement new integrations for the Office of

Risk Management (ORM) and its ERisk system. The ERisk system helps the District file and track incidents related to the safety of District employees and loss of or damage to District property. New data integrations from other systems and datasets such as 311, fleet information, and District buildings provide enhanced analysis and a better understanding of risk management for the District.

The OCTO Middleware team performed over 3.5 million transactions for District systems in 2019.

Citywide Domain/Lookup Tables

Technical Capability: Data Governance Data Quality

The Data Policy designates the CDO with the responsibility of defining Districtwide domain tables and promoting the use of standardized data values and elements across the District's IT enterprise. Improved standardization/quality of data will have many benefits to analysts across the District government, and ultimate users of opendata.dc.gov will benefit as well. For example, designating a table as containing the official list of agency names and abbreviations and promote use of the table by other systems. The following are the <u>designated Districtwide domain tables¹³</u> (also known as lookup tables). Like all DC Open Data, JSON and REST APIs are provided to ease integration with new and existing systems.

2019 Defined Domain Tables

- Address
- Advisory Neighborhood Commission
 name
- Boards and commission list and name
- Charter school name
- Quadrant name
- District agency list and name
- Fire alarm district name
- Fire station name
- Library name

- Square suffix lot list
- Police district name
- Police service area name
- Police sector name
- Public school name
- Single member district name
- Property tax code
- Property use code
- Ward name
- Zoning code

Esri Enterprise Geographic Information System (GIS) Award

Technical Capabilities: Mapping and Analysis



Representatives from DC Health, DC Water, DDOT, DOEE, DPW, HSEMA, OCTO and Planning accept the award from Esri President Jack Dangermond in front of approximately 15,000 GIS Users on July 8, 2019, at the San Diego Convention Center. Photo Credit: Anthony Puzzo, Esri.

Esri Inc., the world's largest provider of mappingrelated software, recognized the District of Columbia Government in July 2019 for **"creating a leading model for governments and businesses throughout the world."** Esri cited the District's use of GIS in "creating greater efficiency, collaboration across departments, transparent and open government, and citizen engagement." The "central GIS/IT infrastructure" provided to all agencies by OCTO and the 1,200 datasets in its open data portal to share necessary information across departments"¹² was also noted.

Empowering DC Employees with Tools and Data

Technical Capabilities: ArcGIS, Tableau, MicroStrategy

OCTO's goal is to remove as many barriers as possible for DC agencies and employees that want to be data driven in pursuing their missions. Through its Data Visualization and Analysis Team, OCTO has software licensing and training in place to make it easy for DC agencies to use modern data analyses and visualization tools.



OCTO's Tim Harwood teaches employees from across District Government how to use the Tableau business intelligence platform.

In 2019, OCTO trained 565 DC government employees and contractors. Training is offered in Esri GIS software, as well as MicroStrategy and Tableau business intelligence software. In 2019 we added classes in the new ArcGIS Pro software and an intermediate class for MicroStrategy. Our classes for DC employees are free; the registration process can be found on the DC intranet.

Planimetric and Orthoimage Data

Technical Capability: Foundational Datasets That Everybody Needs



This image by OCTO's Brian Putz, GIS Analyst, shows new, updated, and removed buildings between April 2017 and April 2019. Areas of change include the Wharf, the Capital River Front, and Buzzard's Point.

OCTO's DC Geographic Information System (GIS) team published updated 2019 planimetric and orthoimages. Publication occurs every other year and employs a Creative Commons Attribution 4.0 license. That means the datasets are open for public use with DC Government credited as the source. The new 2019 data and the archived 2017 data are both available on **opendata.dc.gov**¹⁴.

Where Did My 311 Request Go?

Technical Capability: Data Cards

The Data Development Team continues to invest in the Data Cards User Interface that it first rolled out in 2018 with CrimeCards. dc.gov. In 2019 we improved many of the backend functions that leverage Elasticsearch and launched **311Cards.dc.gov.**

want to explore 10 popular req	ulasts over th		nthe in Ward 5		
	uests over th	le past s mo			
xplore these 6,167 incident	s in the map,	charts, and	tables below or Download this Data		
otals			Мар	In Summary	
	Previous Period	Search Dates	they they are the read	All service requests wen during the past 3 months	s when compared to the
op 10 Service Requests	6,135	6,167 🛧	Bigenetice Bigenetice	same dates a year before	e
Bulk Collection	1,906	1,992 🛧	forest say	un l	
Parking Enforcement	1,072	1,360 🛧	and a standard and a	Your search	Compared to
Parking Meter Repair	58	206 🛧	and a second sec		-
Roadway Signs	258	290 🛧	Nocelay Calumbia A Canady Canady	12/9/2019 - 3/9/2020	12/9/2018 - 3/9/201
Emergency No-parking /erification	246	371 🛧	Total Tota	 12:00 am All service requests 	 12:00 am All service requests
Streetlight Repair nvestigation	335	233 🕹	Menorgane 0.C	e in Ward 5	9 in Ward 5
Trash Collection - Missed	651	643 🕹	Center Ha		
Pothole	834	245 🕹	Legend	•	
Sanitation Enforcement	373	454 🛧			
Residential Parking Permit Violation	402	373 🕹			
Lill Chart ?Help		\simeq	Lill Chart ?Help	Ltdl Chart ?Help	
	ems 1 to 5 of i	6,167 Total		Parking Meter Repair	●1,992 ●1,360
1 2 3 4 5 Ite Bulk Collection Ib: 20=00084425 Address: 208 Q Street NW Vad: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM ©Roadway Signs PM PM	ems 1 to 5 of (6,167 Total		Parking Enforcement Parking Meter Repair Roadway Signs Emergency No-parking V Streetught Repair Investi Trash Collecting Missee Pothol 245 Sanitation Exforcement	erification gation d
Bulk Collection ID: 20-00084425 Address: 208 Q Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM @Roadway Signs ID: 20-00084422	rms 1 to 5 of I	6,167 Total		Parking Enforcement Parking Weter Repair Roadway Signs Emergen 33 No-parking V Streetuat: Repair Investi Trash Collecting 3 Missee Pothol 245 Sanitation Enforcement Residential Parking Perm	erification gation d
Bulk Collection Ib: 20-00084425 Address: 208 Q Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM Roadway Signs	ems 1 to 5 of t	6,167 Total	Chart shows only complete weeks in the chosen date range	Parking Enforcement Parking Meter Repair Roadway Signs Emergency No-parking V Streetught Repair Investi Trash Collecting Missee Pothol 245 Sanitation Exforcement	erification gation d
Bulk Collection Ib: 20=00034425 Address: 208 0 Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM @Radway Signs Ib: 20=00034422 Address: 777 17th Street NE	ems 1 to 5 of t	6,167 Total	Chart shows only complete weeks in the chosen	Parking Enforcement Parking Weter Repair Roadway Signs Emerging Y No-parking V Street Lake Repair Investi Trash Collecting 3 Missee Pothole 245 Sanitation Enforcement Residential Parking Perm Mit Chart 2 Help	erification gation d
Bulk Collection ID: 20=00034425 Address: 208 0 Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM © Roadway Signs ID: 20=00084422 Address: 777 17th Street NE Ward: 5 Zip: 20002		6,167 Total	Chart shows only complete weeks in the chosen date range	Parking Enforcement Parking Meter Repair Roadway Signs Emerging Non-parking V Street Load: Repair Investi Trash Collecting 3 Missed Pothole 45 Sanitation Epforcement Residentian Parking Perm	or 1,360 erification gation it Violation own here? ccidents?
Bulk Collection ID: 20-00034425 Address: 208 Q Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM Roadway Signs ID: 20-00034422 Address: 777 17th Street NE Ward: 5 Zip: 20002 Ordered: 3/8/20, 10:16 PM Parking Enforcement ID: 20-00034412 Address: 230 Adams Street N Ward: 5 Zip: 20002	Έ	6,167 Total	Chart shows only complete weeks in the chosen date range	Parking Enforcement Parking Meter Repair Roadway Silons Emergency No-parking V Streetukat Repair Investi Trash Collecting 3 Sanitation Enforcement Residential Parking Perm Edit Chart PHelp FAQ What is 311 Cards? What 311 requests are sh Why can't I see today's in How is the "previous perin	erification gation d it Violation own here? ucidents?
Bulk Collection ID: 20=00084425 Address: 208 0 Street NW Ward: 5 Zip: 20001 Ordered: 3/8/20, 10:20 PM Roadway Signs ID: 20=00084422 Address: 777 717h Street NE Ward: 5 Zip: 20002 Ordered: 3/8/20, 10:16 PM Parking Enforcement ID: 20=00084416 Address: 230 Adams Street N Ward: 5 Zip: 20002 Ordered: 3/8/20, 10:02 PM Bulk Collection ID: 20=00084379 Address: 1806 Benning Road Ward: 5 Zip: 20002 Ordered: 3/8/20, 8:13 PM	Έ	6,167 Total	Chart shows only complete weeks in the chosen date range	Parking Enforcement Parking Meter Repair Readway Signs Immergency, No-parking V Streetugs Repair Investi Trash Collecting 3 Missee Pothol 9.45 Sanitation Enforcement Residential Parking Perm Idl Chart ?Help FAQ What is 311 Cards? What 311 requests are sho Why can't I see today's in How is the "previous perio for trend cards?	erification gation d it Violation
Bulk Collection Bulk Collection Bulk Collection C	IE NE	6,167 Total	Chart shows only complete weeks in the chosen date range	Parking Enforcement Parking Meter Repair Roadway Silons Imergency No-parking V Streetung Repair Investi Trash Collecting 3 Sanitation Enforcement Residential Parking Perm Idl Chart PHelp FAQ What is 311 Cards? What is 311 requests are sh Why can't I see today's in How is the "previous perio for trend cards? About 311 Cards was created by	erification gation d it Violation own here? ccidents? od" date range selected the the Office of the Ch

311Cards.dc.gov showing the 10 most popular service request types over the past 3 months in Ward 5.



Supporting the 2020 Census

Technical Capability: Address Verification

Not surprisingly, the OCTO Data Team loves the Census. The 2020 Census will open within days of the publication of this report, and we are working in a variety of ways to support Mayor Bowser and her Complete Count Committee. Providing the Census Bureau with the residential address information it needs to make sure every housing unit receives Census information and, if necessary, an enumerator is a key role of the OCTO Data Team. The DC Office of Planning, working with the support of the DC GIS Team, was able to add over 36,000 housing units to Census records, which speaks to the growth of the city since the 2010 Census.

Good housing records are essential to understanding how much progress we are making toward achieving Mayor Bowser's goal of 36,000 new homes in the District by 2025.



For more information on the 2020 Census, visit dccensus2020.dc.gov



What's Going on With Our Kids?

Technical Capability: Data Dashboards

In the 2019 CDO report we announced that OCTO's two supported business intelligence systems, Tableau and MicroStrategy, had new public-facing servers so agencies can more easily publish dashboards and allow the public to interactively explore data. In response to this, here are leading examples of how DC agencies have used this new capability to share information with the public:

- Child and Family Services Administration, Dashboard: cfsadashboard.dc.gov
- Deputy Mayor for Education, EdScape: edscape.dc.gov





How many students commute out of Ward 7 to attend school and where do they go?

What is the breakdown of children in foster care by race?



Landing Businesses in DC's Opportunity Zones

Technical Capability: Community Engagement Sites

Traditionally, the OCTO Data Team has specialized in pushing data out to the public. Opendata. dc.gov, Web mapping, apps, and dashboards all make the District's data available and help the public understand and query that data. Today we are attempting to move beyond that by leveraging data through two-way participation with stakeholders–public, internal, or both. District agencies are using ArcGIS Hub Initiative sites to organize and create micro-open data sites of their own. DC's ArcGIS Hub packages Open Data DC, ArcGIS Online web maps, and engagement apps like surveys. These sites are driven by a focused agency program or activity designed to complement dc.gov sites. Typically, our sites allow residents and public users to choose the level of engagement they are comfortable with. For example, the Deputy Mayor for Planning and Economic Development worked with OCTO to develop the DC Opportunity Zone Marketplace (ozmarketplace.dc.gov). The site supports a variety of levels of engagement from browsing all the way to volunteering. These are not new capabilities for the government, but it is an expansion of what the Data Team can help agencies accomplish. Therefore, it is imperative that they be used in close coordination with the OCTO Web Team and agency communications and engagement teams



The Opportunity Zones program provides tax incentives for equity investments in businesses and commercial projects in designated Census tracts. Users can post qualified investment opportunities. New businesses can request assistance from DC-based professional service providers.

Transportation

What Are All Those Scooters Doing?

Technical Capabilities: Big Data, Distributed Analysis, Streaming Data

Agencies have begun to use the DC Data Lake that was first described in the 2019 CDO report. For example, The District Department of Transportation (DDOT) began working with dockless bikeshare providers in 2017, and by 2018 the program had expanded to include the very popular e-scooters. DDOT's relationship with the dockless providers allows the government to receive information about the scooters' movements so important questions can be asked and answered. For example, to what extent are:

- Scooter trips complementing public transit or replacing public transit trips?
- Scooter riders making use of facilities, such as bike lanes, that enhance their safety, or are they just taking the shortest path?
- Scooter companies servicing all areas the city equitably?

The DC Data Lake is helping DDOT answer these questions by enabling the ability to ingest and process analysis a large volume of scooter data. DDOT, with OCTO, is successfully receiving data from every dockless vehicle whose provider has been awarded a permit. Currently we are receiving and storing approximately 30,000 anonymized dockless locations daily, and more providers may have licenses by the end of March. Moreover, we can securely expose that data to authorized users with standard analysis tools such as Esri, Microstrategy, and Tableau as well as data science tools such as Jupyter Notebooks. DDOT is already implementing policies based on their analysis. Does the dockless map to the upper right look equitable? DDOT doesn't think so, therefore is requiring the new 2020 dockless permit holders to do more. You can see what DDOT is doing by looking at the Dockless Equity Emphasis Areas as shown in the map below on the Open Data Portal.





Public Safety Portal Expanded

Technical Capabilities: Streaming Live Data, Securing Sensitive Data

Police Department (MPD) and the Homeland Security and Emergency Management Agency (HSEMA). MPD's Investigative Support Section (ISS) analysts rely on the ISS App that allows them to support investigations in the District. HSEMA continues its investment in developing content to promote situational awareness for its operators and emergency operations center's staff.

During the 2019 World Series, several dashboards and applications were used that displayed realtime feeds including police and fire automated vehicle location/radio locations, computeraided dispatch events, traffic data, and mass transportation data. One real-time data feed created for the World Series event was the WMATA Rail Ridership feed. This feed is used to display exit and entrance data for each of the WMATA metro stations, providing end users a real-time look into how many people were using metro services in and around Nationals Park. Following these events, OCTO archives this WMATA feed for after-action analysis.





Mosquitos: Map and Eliminate Them

Technical Capability: Field Data Collection

Since most health data is protected under level 3 and 4 classification there are few examples that we can publish but one successful use of field data resources continues to be DC Healths collaboration with other agencies to test mosquito populations for the Zika and West Nile viruses by trapping mosquitos throughout the District. DC Health also partners with the National Park Service and the U.S. Department of Defense to access federal properties. The data collected in the field helped DC Health find and address inefficiencies in the larvacide program.



opendata.dc.gov



What Is the District Government Doing for the Environment?

Technical Capability: Data Storytelling

The tools OCTO provides can be used to tell a story. Environment data stories have been the most popular. For example, in 2019 the Department of General Services produced "**The Story of a Banana Peel.**" Students can trace the story of their lunchtime banana from a farm in Ecuador through the Port of Philadelphia to their plate in a DC Public School cafeteria. They can then learn how the peel is composted and becomes fertile dirt that is used at Nationals Park and elsewhere. There are many such environmental stories on the Open Data Portal.



Legislative Affairs, Religious Affairs, Women's Policy and Initiatives. The Deputy Mayor for Greater Economic Opportunity was dropped.

8. opendata.dc.gov/datasets/a4968111fffc423ea3a83c51a4242e66_22

14. <u>opendata.dc.gov/search?q=planimetric&sort=-modified</u>

^{4. &}lt;u>opendata.dc.gov/datasets/enterprise-dataset-inventory</u>

^{5.} Consolidated with EOM for purposes of the inventory were the offices for Asian and Pacific Islanders Affairs, Federal and Regional Affairs, Policy and

^{6. &}lt;u>opendata.dc.gov</u>

^{7.} opendata.dc.gov/datasets/enterprise-dataset-inventory-level-0-not-on-open-data-dc

^{9.} os.dc.gov/page/annual-reports

^{10. &}lt;u>opendata.dc.gov/datasets/cumulative-foia-report-for-fy-2019</u>

 $^{11. \} may or .dc. gov/release/washington-dc-recognized-bloomberg-philanthropies-use-data-improve-delivery-government$

esri.com/about/newsroom/announcements/esri-honors-organizations-for-exceptional-use-of-geospatial-technology
 dcgis.maps.arcgis.com/apps/MinimalGallery/index.html?appid=a7271bd43ec345ed97d6374e71075e07

Where We are Going: The Draft Tech Plan and New Goals

As discussed in the introduction, this is the third CDO annual report, and it is the first year in which this document has been produced within the context of an overriding Draft Technology Strategic Plan for DC: Unleashing the Possible. Today, with this new strategic direction, we are empowered and looking forward to using data to solve the complex problems facing DC Government.

It is expected that the Draft Tech Plan will be edited and updated including edit to Commitment 4 which embraces the data goals and strategy. At the time of publication of this report, March 10, 2020, feedback is still being collected at **techplan.dc.gov.**

The following are 2019 goals and progress, as well as, the new initiatives and recommended tasks the OCTO Data Team and it's partners are working on to achieve Commitment 4 of the Draft Tech Plan.

Initiative 4.1

Identify high-value mission use cases to improve the use of data in service of specific Mayoral initiatives and focus areas.

2019 Goal: Train agency communications and engagement teams in data storytelling

As Mayor Bowser has demonstrated with her mayor. dc.gov website, opendata. dc.gov offers many ways to explain and track policies using open data. Familiarizing communications teams with the capabilities should drive publication and utilization of open data.

Some Progress

A data team representative speaks to every web training class hosted by the OCTO Web Team. However, Comms teams still need more exposure to what the OCTO Data Team can help them achieve:

- Interactive "Hub" sites, like the one developed by DDOT's Urban Forestry Administration
- Public facing dashboards like EdScape developed by the Deputy Mayor for Education
- Add open data integrations like mayor.dc.gov, which was featured in last year's report.

Initiative 4.1 Continued

Identify high-value mission use cases to improve the use of data in service of specific Mayoral initiatives and focus areas.

New Goal

Support Mayor Bowser's housing focus area under the Accountability Report.

New Goal Focus on one DC government cluster in depth. We cited our work in support of the Office of Planning for the U.S. Census and our contribution to Mayor Bowser's focus area on housing. But we can be doing more to support housing agencies. Next year's report shall feature an "accomplishment" with a housing agency.

While continuing to support <u>all</u> DC agencies, next year we will work closely and voluntarily with one DC government cluster (Education, Health and Human Services, Internal Services, Operations and Infrastructure, Planning and Economic Development, or Public Safety and Justice) to go more in-depth in understanding the cluster's data and supporting its mayoral initiatives. The 2021 CDO report shall include a section that focuses on the sccomplishments of these efforts.

Initiative 4.2 Streamline policy and process to share and use data.

2019 Goal: (*Revised from 2018*) **Develop eMOU system to support data-sharing agreements**

Not all data can be open, and the Data Policy calls on the CDO to develop a "streamlined process for interagency data sharing." A data-sharing agreement is a document of agreement between two agencies in which the data steward agrees to share specific data with another agency subject to certain terms and limitations. Faster execution, with better tracking and enforcement of datasharing agreements, is needed across the District government. OCTO already maintains a system known as "eMOU," where MOU stands for memorandum of understanding. With modifications, such as multilateral agreements, the eMOU can be adapted to handle standardized data-sharing agreements.

• The tool will need to support multilateral or brokered agreements, not just agreements between two agencies.

Progress but not Enough

We now have a draft multilateral data-sharing agreement template that establishes a "facilitator" role for OCTO. The template has recently been turned over for legal review and then automation can begin.

Initiative 4.2 *continued* Streamline policy and process to share and use data.

2019 Goal: Move open datasets to the Open Data Portal

OCTO will work with agencies, the Open Government Advisory Group, and the community to prioritize posting the remaining 216 open enterprise datasets identified in the EDI that are not yet on the open data portal.

2019 Goal: Develop API best practices

Part of the CDO mission is to facilitate data sharing. APIs are a common tool for fast, flexible, and secure data integration. OCTO maintains minimal API standards for District agencies on its intranet site, but much more guidance is required. This is especially true where more and more, but not all, data is moving to cloud-based software and service applications.

2019 Goal: Develop cutting-edge data platform to support analytics

Now is the time to push utilization of the Data Lake.

- The Lab @ DC has agreed to migrate at least one agency's datasets into OCTO's searchable "Data Lake" and undertake at least one multiagency Data Science Project with the system.
- The OCTO Data Team will gather and catalog as much Level 1 and Level 2 data as we can to populate data.in.dc.gov.

Progress

Per the EDI, we have 51 more datasets on opendata.dc.gov than we had last year. That said, the inventory now contains 245 classified as Level 0 that are not on the open data portal. This will be an ongoing task.

Progress

The Data Team has prepared a draft "API and Middleware Policy," which is going through the review process. We will continue to work on best practices in the coming year.

Progress

The DC Data Lake is up and running. This year we focused on utilization and made progress with three significant customers:

- DDOT is using the lake to store and analyze dockless mobility data.
- OCTO's DC Net program, as part of a proof of concept for Low-Power Wide Area Network, is bringing in smart city sensor data.
- The Lab @ DC has approved the data lake for some of its sensitive datasets.

The less sensitive Level 1 and Level 0 data have been loaded in the lake.

Initiative 4.2 continued

Streamline policy and process to share and use data.

2019 Goal: Move open datasets to the Open Data Portal

The District is still failing to track a significant percentage of FOIA requests inside its enterprise system. Therefore, we are resolved to

- Continue regular training of FOIA Officers.
- Identify agencies and FOIA Officers not using the system, and work with them to increase system utilization.
- Begin publishing the District's Annual FOIA Reports as open data in addition to PDF format so that the reports can be more easily analyzed.
- Work with agencies to review and publish the priority list from the FOIA section of this report.

Also, on behalf of FOIA requesters, we will review and improve or replace the FOIAXpress Public Access Link (PAL).

2019 Goal: Provide more assistance to agencies to comply with FOIA

With opendata.dc.gov, users can link to filtered data. The OCTO Data Team trains web editors on this technique, so they can link to agency-specific records from agency FOIA and open government web pages.

Accomplished

Per the revised 2019 goal:

- We did continue training FOIA Officers and initiated a new two-day class.
- We have identified agencies not using the system in this report.
- We have published the District's Annual FIOA report as open data.

Progress

Our vendor for FOIAXpress, AINs, has made significant progress on a mobile-friendly version of the public access link. We hope to test it this spring and release it soon after.

Not much progress

Publishing data from priority section of last year's report.

On Hold

A larger redesign of DC.gov is in the works. Improvements to our open data search and data curation are also required to allow this to work well.

Initiative 4.2 continued

Streamline policy and process to share and use data.

New Goal:

Develop technology and processes to help agencies with data anonymization.

There are a variety of ways to summarize or mask PII to make it harder to identify individuals within a dataset. No anonymization is full proof, but OCTO should offer standard practices and tools that make easier to anonymize Level 3 for sharing and analysis. OCTO already has some capital funds dedicated to this project.

New Goal: Leverage the EDI to promote data sharing Develop an interactive dashboard to allow the public and agencies to explore the EDI. This will improve openness and generate more requests to share data.

Initiative 4.3

Support agency use of big data and artificial intelligence

2019 Goal: Develop cutting-edge data platform to support analytics

Now is the time to push utilization of the Data Lake.

- The Lab @ DC has agreed to migrate at least one agency's datasets into OCTO's searchable "Data Lake" and undertake at least one multiagency Data Science Project with the system.
- The OCTO Data Team will gather and catalog as much Level 1 and Level 2 data as we can to populate data.in.dc.gov.

Progress

The DC Data Lake is up and running. This year we focused on utilization and made progress with three significant customers:

- DDOT is using the lake to store and analyze dockless mobility data.
- OCTO's DC Net program, as part of a proof of concept for Low-Power Wide Area Network, is bringing in smart city sensor data.
- The Lab @ DC has approved the data lake for some of its sensitive datasets.

The less sensitive Level 1 and Level 0 data have been loaded in the lake.

Initiative 4.3 *continued* Support agency use of big data and artificial intelligence

New Goal Develop a plan for leveraging Artificial Intelligence (AI) The District agencies and OCTO don't currently have a shared understanding of what AI is or how we would apply it in support of the Mayor's Focus Areas. Together with SmarterDC we shall develop and publish such an understanding and direction.

New Goal Revitalize the DC Data Science User Group OCTO helped to found a Data Scientist user group for District employees. The group promoted information sharing and pest practices for data science projects and predictive analytics. Meetings have fallen off due to staff turnover. The group should be revitalized in the coming year with regular meetings and increased membership.

Conclusion

The CDO and the OCTO Data Program strives to help agencies manage and use data so they can be more efficient and effective. Although agencies are our direct customers, the underlying reason for everything we do is to serve the public. We want to hear about the datasets you need and the ideas you have. What did you like about this report and inventory list? What could we be doing better? What are we not doing at all that we should be doing? Please reach out by email to **open.data@dc.gov** or Twitter **@opendatadc** and let us know what you think and how we can do better.