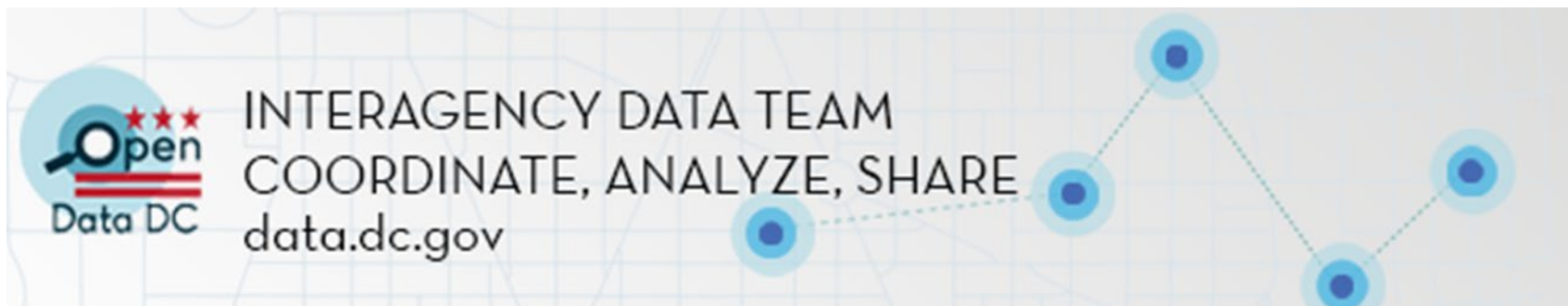


# Interagency Data Team

March 13, 2019



# Welcome, agenda Data team news

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

# Agenda



## ***Welcome, News & Updates***

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



## ***Enterprise Dataset Inventory Analysis***

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



## ***Chief Data Officer's Annual Report***

Barney Krucoff, Chief Data Officer  
Office of the Chief Technology Officer



## ***MicroStrategy World – New Features & Solutions***

Bircan Tekin, BI Solutions Architect  
Office of the Chief Technology Officer

# Business Intelligence - Tableau Updates

Office of the Chief Technology Officer

**Tableau  
Internal  
Production @  
version 2018.1.1**

Hold off on  
Desktop  
upgrades (keep  
at 2018.1.x if  
publishing to  
the server)

**Tableau  
Internal  
Failover  
updated to  
version 2018.3.**

Announcement  
planned  
failover once  
testing is  
complete

**Tableau Public  
Production @  
version 2018.1.1**

DDOT –  
[Circulator](#) and  
[MPD crash data](#)

EOM – [By The  
Numbers](#)

OP – [SDC Portal](#)

**Tableau Site  
Administrators  
quarterly  
meeting  
scheduled for  
4-3-2019**

Invitations will  
go shortly

DC Government

# Business Intelligence - MicroStrategy Updates

Office of the Chief Technology Officer

## MicroStrategy Production @ version 10.11

New physical servers expected FY19 end of Q3.

Webserver OS migrations - performance issue found during testing with Dossier.

## Lower MicroStrategy Environments

Enterprise/Self Service

Development

Test

QA

## MicroStrategy Public

Pilot with DBH

Guidelines for imbedding analytics in Agency websites and SharePoint developed.

## MicroStrategy Stake Holders Meeting

Will be scheduled spring 2019

DC Government

# Geographic Information Systems - DC GIS

Office of the Chief Technology Officer

**ArcGIS Server  
Production @  
version 10.6.0**

10.6.1  
deployment  
notice in 2-3  
weeks.

ArcGIS server  
with Roads and  
Highway  
extension @  
version 10.6.1

**ArcGIS Portal  
@ version  
10.6.1**

**Streaming  
services**

ESRI GeoEvent  
will be in  
production  
@10.6.1

Apache Kafka  
DSP Streaming  
data (FY 19 Q2  
production)

**Data**

2019 EDI

Pictometry  
winter capture  
complete

Citywide  
NearMap License  
Lapsed

Spring Ortho-  
Planimetric flight  
kicking off

DC Government





# Tableau Training

## Introductory training

March 20 and 21 (Full), April 3 and 4 (Full)

May 2 and May 3, May 8 and May 9

## Intermediate training

March 26 and 27 (two seats available)

April 10 and 11 (full)



<https://octo.in.dc.gov/service/tableau-training-courses>



# MicroStrategy Training

Introductory training

April 30 and May 1 (10 seats left)

May 28 and 29



<https://octo.in.dc.gov/service/microstrategy-introductory-training>





# GIS Training

230 Overview DC GIS Services: April 10<sup>th</sup>, May 8<sup>th</sup>, September 11<sup>th</sup>

233 ArcGIS Online Virtual Sessions (scheduled in the near future.)

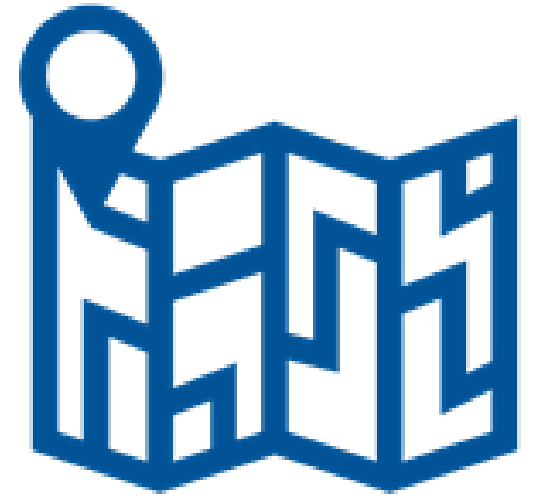
ArcGIS Pro training is under development



<https://octo.in.dc.gov/service/enroll-gis-training>

# GIS / Data Science User Meeting

- GIS – March 21<sup>st</sup> (Via WebEx) contact [Eva.Reid@dc.gov](mailto:Eva.Reid@dc.gov)
- Data Science @ DC – March 13<sup>th</sup> 4-5pm (today!) @ 200 I street in OCTO Room 5009 contact [Peter.Casey@dc.gov](mailto:Peter.Casey@dc.gov)



# Enterprise Dataset Inventory Analysis

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

Figure 1  
How many agencies recorded enterprise datasets?

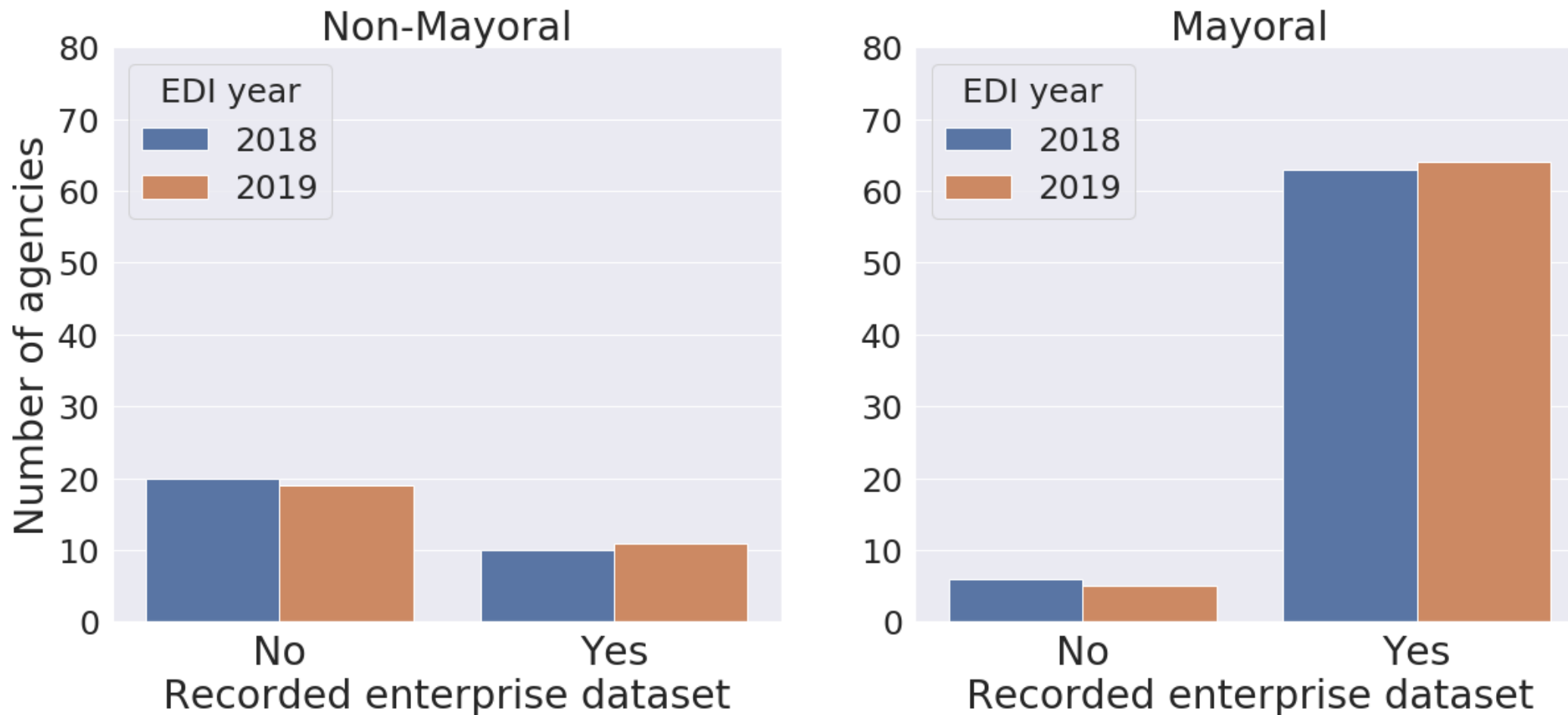


Figure 2  
How many enterprise datasets did agencies record?

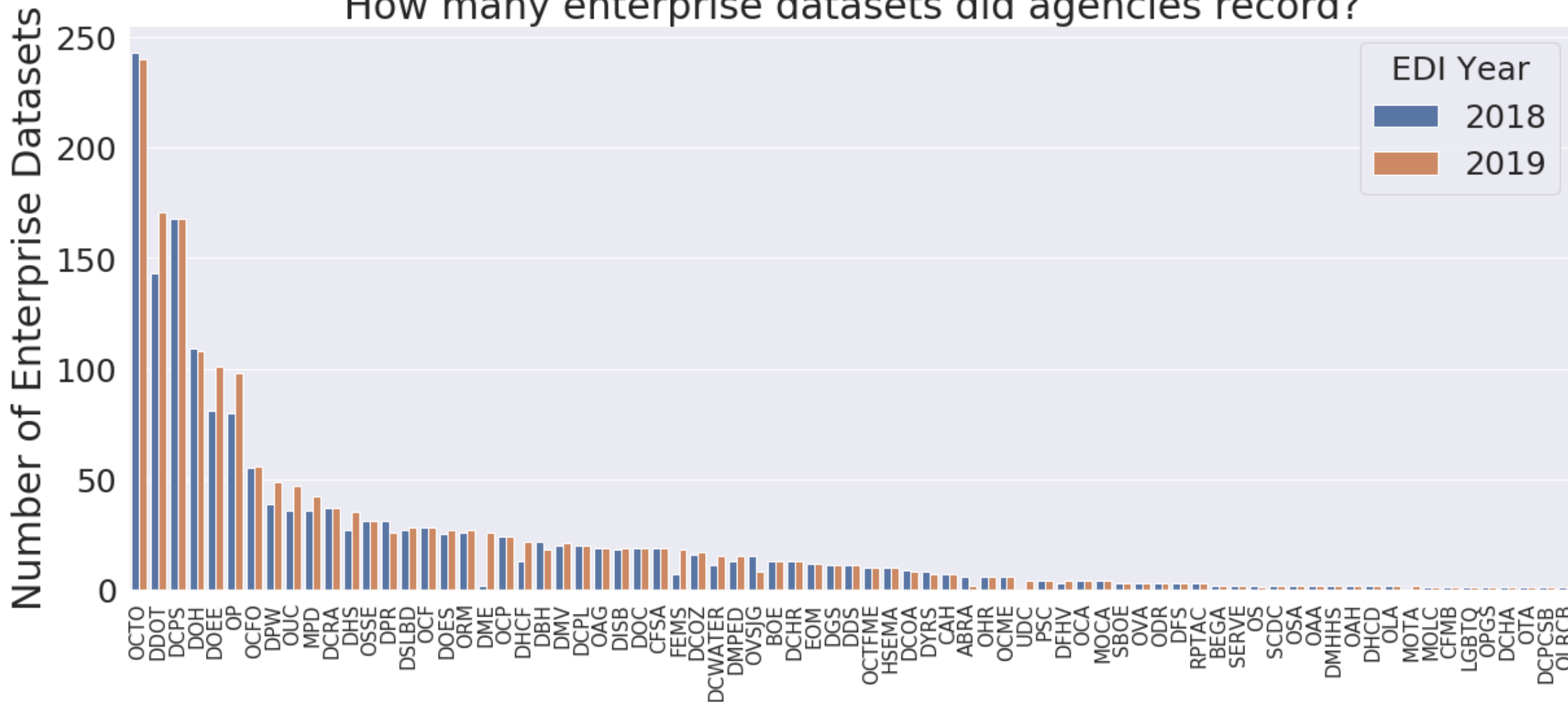
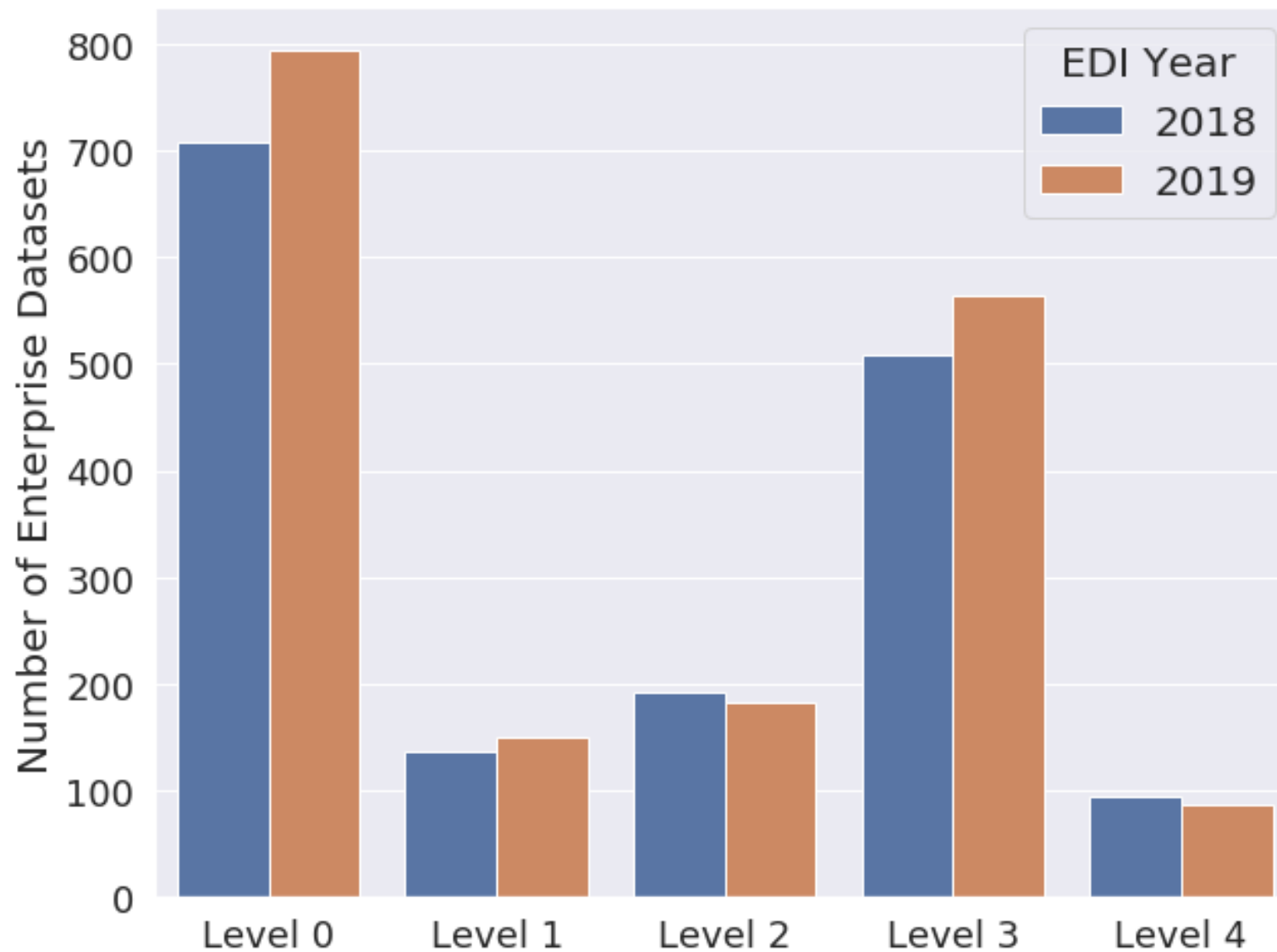


Figure 3  
How were datasets classified?



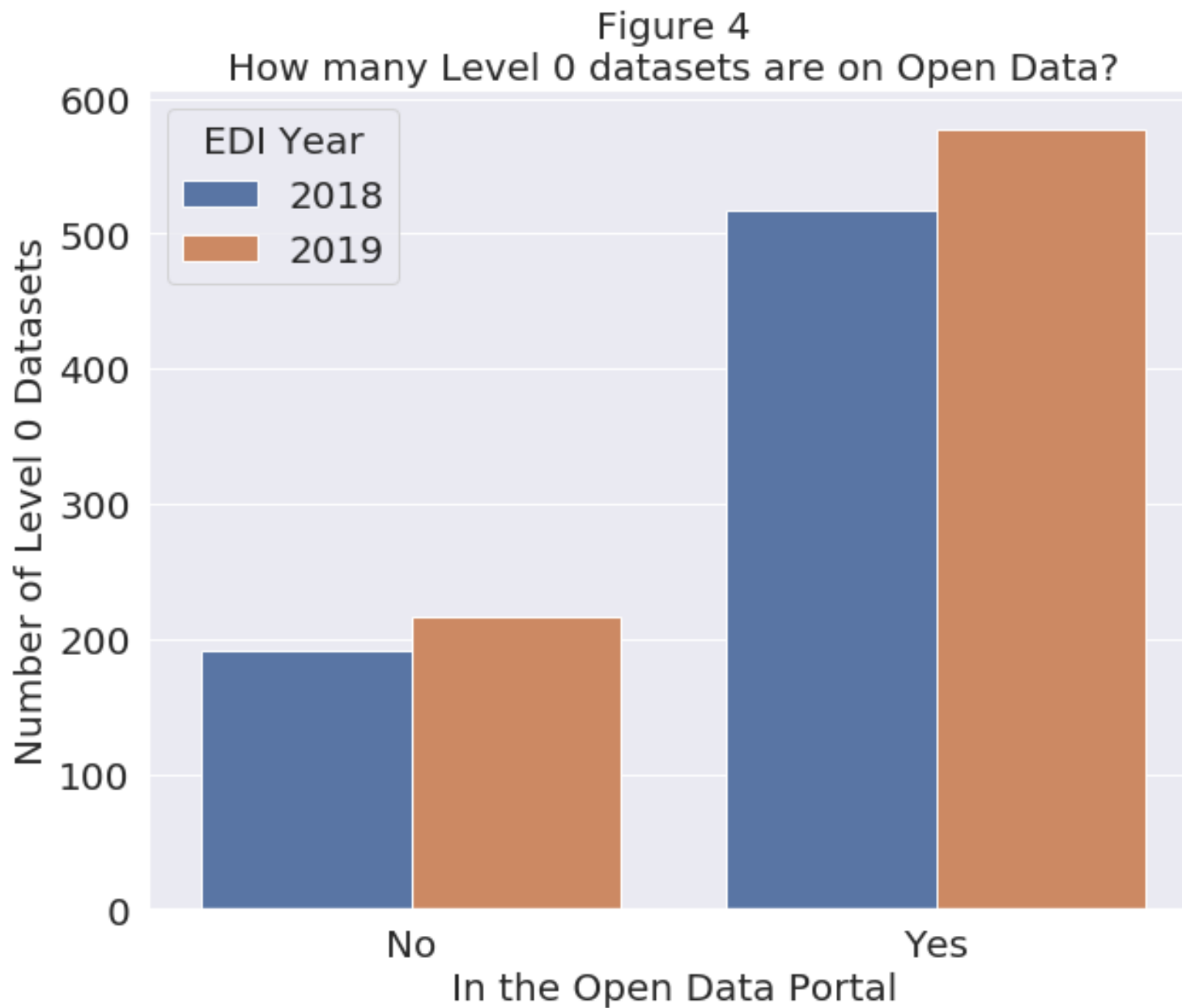
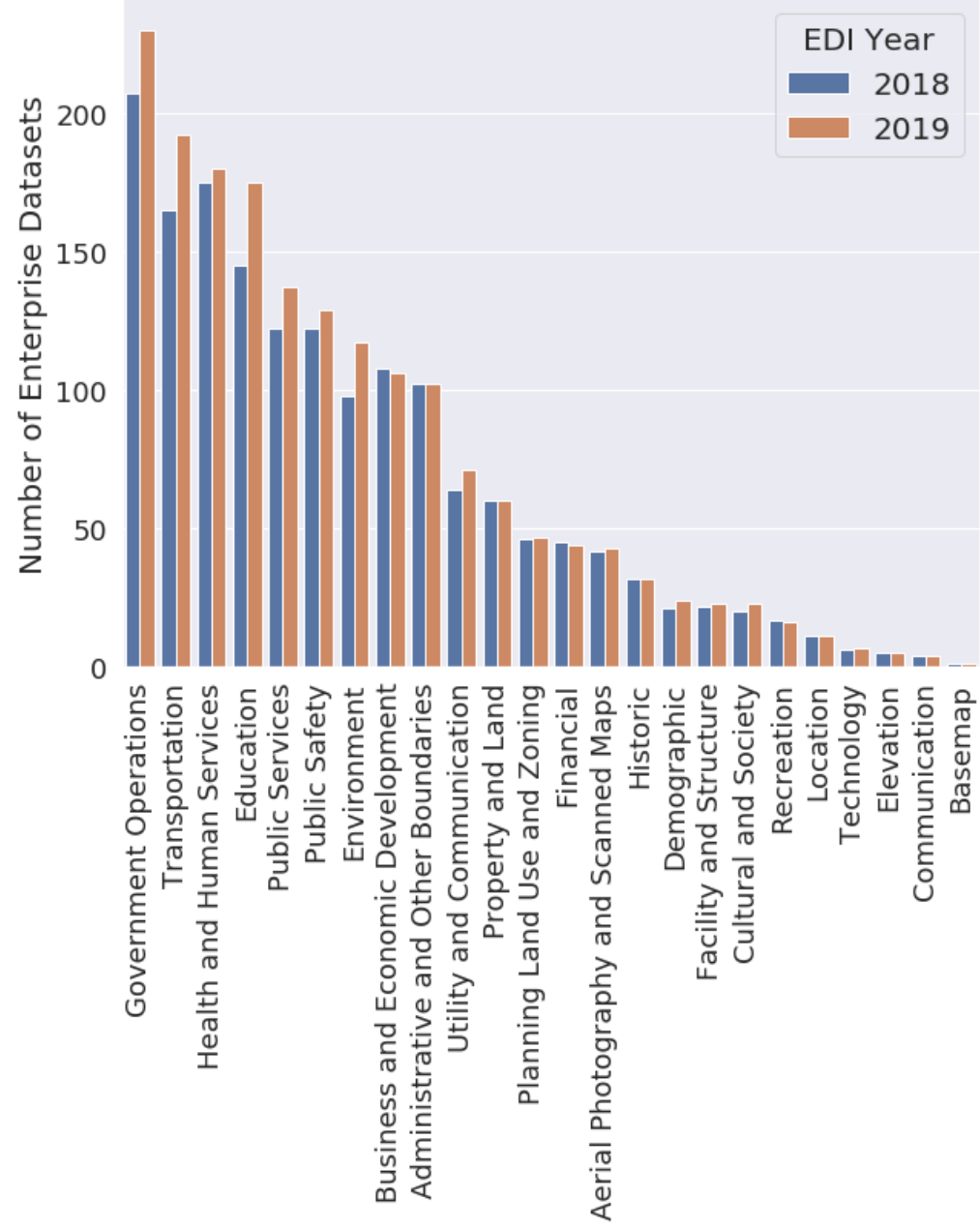


Figure 5  
How were datasets categorized?





# Chief Data Officer's Annual Report

Barney Krucoff, Chief Data Officer  
Office of the Chief Technology Officer

# Chief Data Officer's Annual Report

GOVERNMENT OF THE  
DISTRICT OF COLUMBIA

March 11, 2019



## Report Sections

- EDI Analysis
- FOIA – Open Data Feedback Loop
- Accomplishments
- Accountability

# FOIA and Open Data Should Work Together

The Data Policy envisioned a “nexus between FOIA and level zero open datasets, where FOIA and Open Data are distinct but complementary practices.”

- 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 request-tracking 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.

# Use of FOIAXpress

FOIAXpress utilization measures by calendar year	Calendar year 2017	Calendar year 2018
# of requests and appeals tracked in FOIAXpress	7,346	8,120
# of agencies enrolled in FOIAXpress	53	55
# of agencies processing at least one request in system	51	52
# of current active DC government users	112	141

FOIA Requests	FY 2017	FY 2018	Increase
Total reported by EOM	8,274	10,450	26%
Tracked in FOIAXpress System	6,762	7,943	17%
% in FOIAXpress	82%	76%	

## Requested Not on Open Data

- Arrests MPD
- Condom distribution statistics, DCPS, DOH
- Dockless bikes and scooters, DDOT
- Facility surveys, PCSB, DGS, DCPS
- Health inspections, DOH
- Health professional licenses, DOH
- Moving violations by DC-owned vehicles, DMV
- Property tax credits and penalties, OCFO
- Seizure data, MPD
- Vacant property determinations, DCRA

## Frequently Requested on Open Data

- Building permits, DCRA
- Business licenses, DCRA
- Underground storage tanks and similar, DDOE

OCTO Data Team and Selected Agency  
**Accomplishments**



Theft from Auto

852

Thefts from automobile incidents in the last 30 days.  
Source: Crime Incidents in the Last 30 Days



Public Safety



Education



Environment

Number of Public Trees

158,746

Street trees planted in tree boxes and cared for by DC.  
Source: UFATree2\_FeatureServer

Projects Under Construction

137

Affordable housing projects currently under construction.  
Source: Affordable Housing



Housing

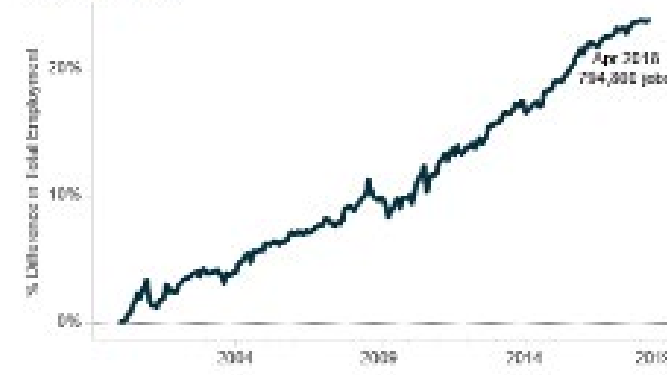


Economic Development

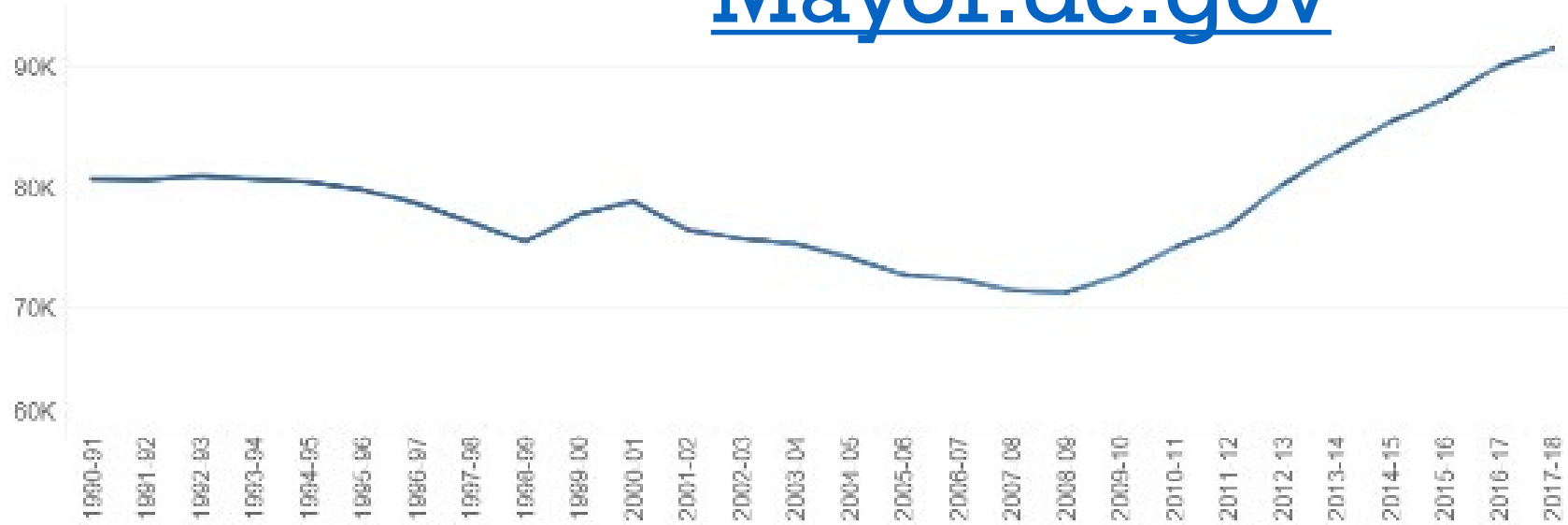


Transportation

Total Employment in DC | Growth Since 2005  
Jan 2003 to Apr 2018

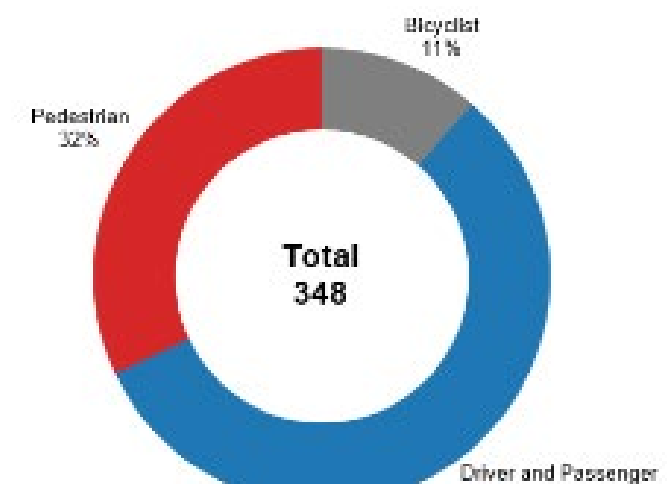


Public School Enrollment  
1990-2017



[Mayor.dc.gov](http://Mayor.dc.gov)

Number of Persons Injured in Crashes by Type, 2018



# Publication DC GIS Achieves



1999

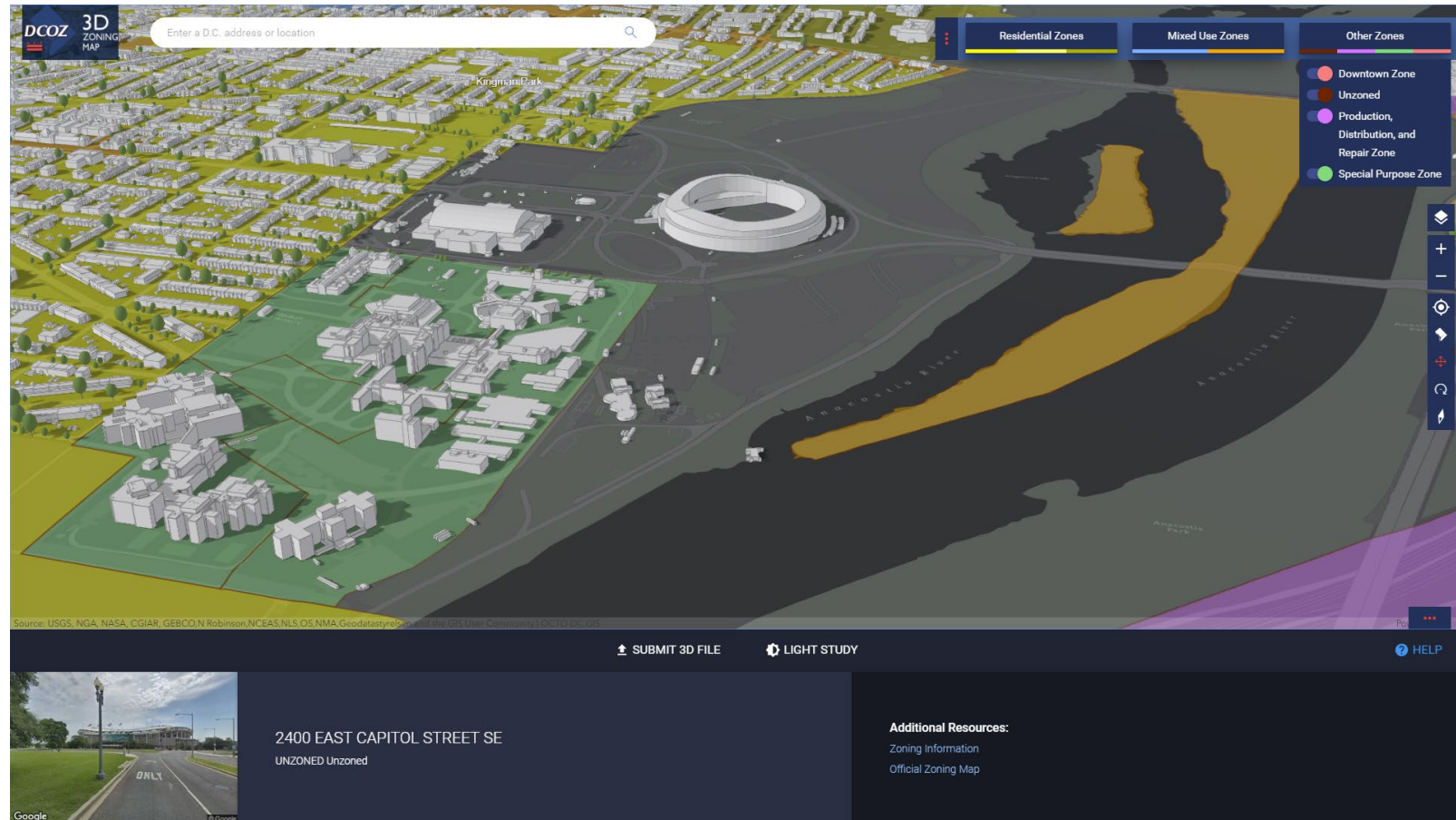
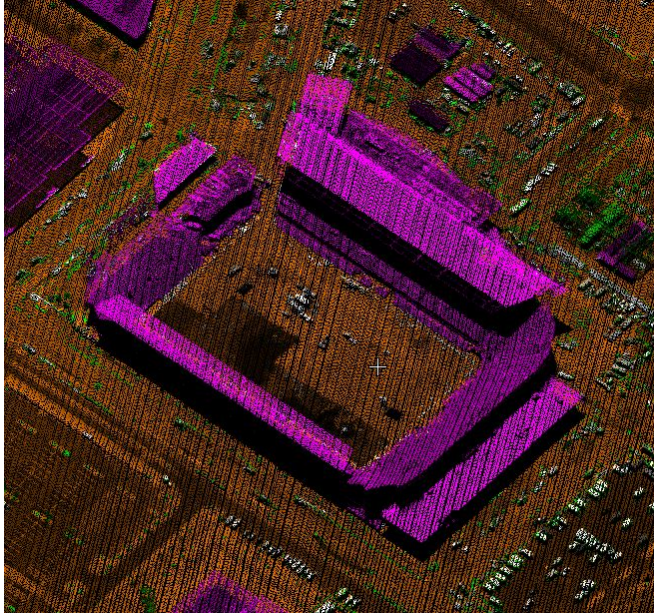
Columbia Heights



2017

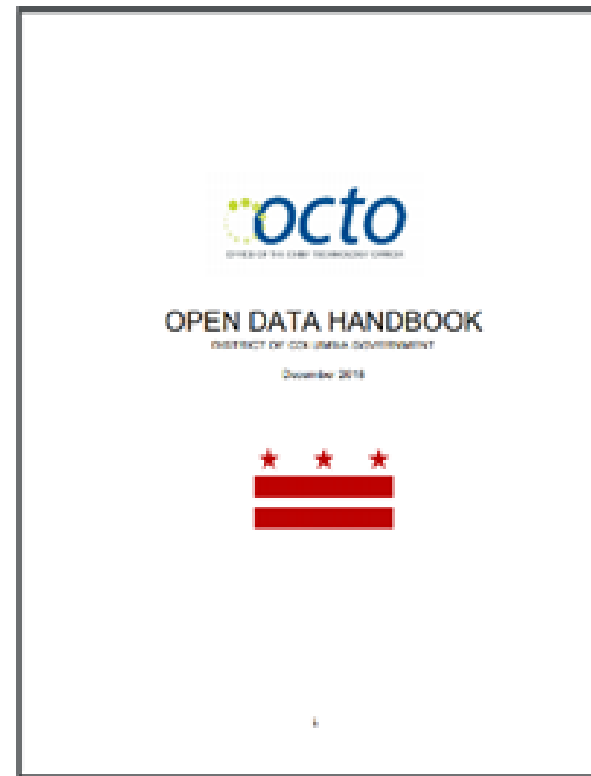


# LiDAR & 3D Zoning



# Open Data Handbook






Better standards



# DC Data Lake



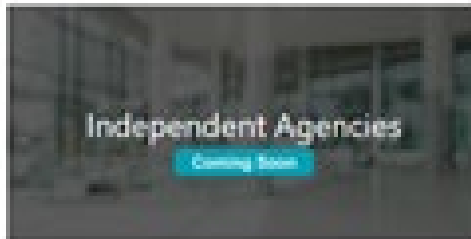
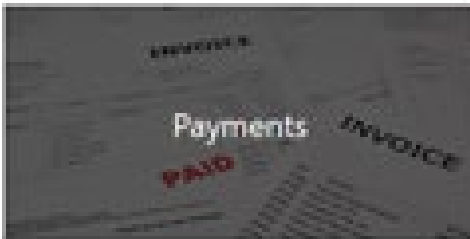
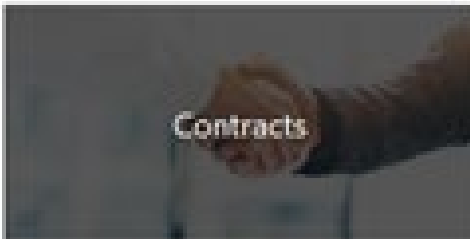
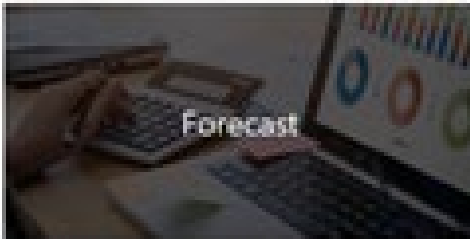
A future meeting will be dedicated to the lake!

<b>Discoverability</b>  	Search Engine for City Staff at <b>data.in.dc.gov</b>
<b>Usability</b>  	Methods and tools to bring the data into City's current data toolbox: MicroStrategy, Tableau, ArcGIS, Excel (csv), <u>Jupyter</u> notebooks, Spark, Kafka
<b>Security</b>  	Governance by Data Policy, encryption and role-based authentication at all phases  
<b>Flexibility</b>  	Unstructured and Structured data in a single platform for Advanced Analytics and AI.

## Access All of DC Government's Business Opportunities Today

Welcome to the Contracts and Procurement Transparency Portal! The Contracts and Procurement Transparency Portal was created by the Office of Contracting and Procurement (OCP), under the leadership of Mayor Muriel Bowser, as a public clearinghouse for all information related to the District Government's contracting and procurement efforts.

What would you like to search today?



# OCP Procurement Transparency

## New and Improved Open Data

- [Contracts](#)
- [Payments](#)
- [Forecast](#)
- [Purchase Orders](#)
- [Solicitations](#)

OCP launched a new Contracts and Procurement Transparency Portal (<https://contracts.ocp.dc.gov>)

# Goals Last Year / This Year

2018 Goals (March 10, 2018 to March 10, 2019)	Accountability and Revised Goals (March 10, 2019 to March 10, 2020)
<b>Move Open Datasets to the Opendata.dc.gov Portal</b> 195 open enterprise datasets identified in the EDI that are not yet on the Open Data Portal.	<b>Partially Accomplished</b> Of those 195 datasets, 43 have been published on opendata.dc.gov. An additional 7 datasets turned out not to be structured <b>Revised 2019 Goal</b> Prioritize these 189 datasets for publication prior to next year's report. The revised list is online.
<b>Develop eMOU System to Support Data Sharing Agreements</b> 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.	<b>Not Started</b> <b>Revised 2019 Goal</b> Facilitating and tracking data-sharing agreements remains a goal for 2019. Some addition considerations: <ul style="list-style-type: none"> <li>• There is considerable legal work still required to come up with acceptable agreement templates.</li> <li>• The tool will need to support multilateral or brokered agreements, not just agreements between two agencies.</li> <li>• EMOU or more of a document management approach.</li> </ul>
<b>Develop and Publish a Data Submission Guide</b>	<b>Accomplished!</b>

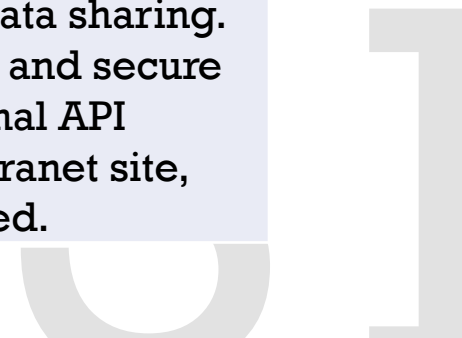


# Goals Last Year / This Year

2018 Goals (March 10, 2018 to March 10, 2019)	Accountability and Revised Goals (March 10, 2019 to March 10, 2020)
<b>Improve FOIA Request Tracking</b>	<p><b>Progress, but Not Enough</b></p> <p><b>Revised 2019 Goal</b></p> <p>The District is still failing to track a significant percentage of FOIA requests inside its enterprise system. Therefore, we are resolved to:</p> <ul style="list-style-type: none"> <li>• continue regular training FOIA Officers;</li> <li>• identify agencies and FOIA officers not using the system, and work with them to increase system utilization;</li> <li>• 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; and</li> <li>• work with agencies to review publish the priority list from the FOIA section of this report.</li> </ul> <p>Also, on behalf of FOIA requesters we will review and improve or replace the FOIAXpress Public Access Link.</p>

# Goals Last Year / This Year

2018 Goals (March 10, 2018 to March 10, 2019)	Accountability and Revised Goals (March 10, 2019 to March 10, 2020)
<b>Develop Cutting-Edge Data Platform to Support Analytics</b> Develop a Hadoop-based big data backend capable of hosting critical data and support while lowering application development and support costs.	<b>Significant Progress</b> Subject of future meeting  <b>Revised 2019 Goal</b> Now is the time to push utilization of the Data Lake. <ul style="list-style-type: none"> <li>• 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.</li> <li>• 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.</li> </ul>
<b>New Goal</b>	<b>Develop Application User Interface (API) Best Practices</b> 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, however much more guidance is required.



# Goals Last Year / This Year

2018 Goals (March 10, 2018 to March 10, 2019)	Accountability and Revised Goals (March 10, 2019 to March 10, 2020)
<b>Provide More Assistance to Agencies to Comply With FOIA for public facing websites must contain specific information, including:</b> <ul style="list-style-type: none"> <li>• public employee salary information;</li> <li>• administrative staff manuals and instructions;</li> <li>• statements of policy;</li> <li>• information dealing with the receipt or expenditure of public funds;</li> <li>• budget information;</li> <li>• minutes of public meetings;</li> <li>• frequently requested public records;</li> <li>• District-wide and agency FOIA reports; and</li> <li>• organizational chart.</li> </ul>	<b>Largely Accomplished</b> <p>OCTO Web Editors have updated all the open government pages that they manage. Some changes depend on agencies.</p> <p>Revised 2019 Goal</p> <p>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.</p>
<b>New Goal</b>	<b>Train Agency Communications and Engagement Teams in Data Storytelling</b> <p>Familiarizing communications teams with the capabilities should drive publication and utilization of open data.</p>



# Recommendations

2018 Recommendations (March 10, 2018 to March 10, 2019)	Status
Move EDI and CDO Report Deadline to the Monday of Sunshine Week	Done!
Adopt EDI Participation as a Key Performance Indicator for All Agencies	Done!
Adopt Reasonable and Uniform Retention Policy for Email	Remains a Recommendation
Encourage Independent Agency Participation in the Enterprise Dataset Inventory	Some progress but More to Do

# MicroStrategy World New Features & Solutions

Bircan Tekin, BI Solutions Architect  
Office of the Chief Technology Officer

# MicroStrategy 2019

- HyperIntelligence
- Federated Analytics
- Natural Language Queries
- New R and Python Packages
- New Out of the Box Visualizations

# HyperIntelligence

- HyperIntelligence that delivers zero-click intelligence to any web page or web application.
- Deploy contextual insight to any website, screen, wall, device or application
- Get answers in real time without interrupting the workflow – no clicks required



HYPERCARD



HYPERVOICE



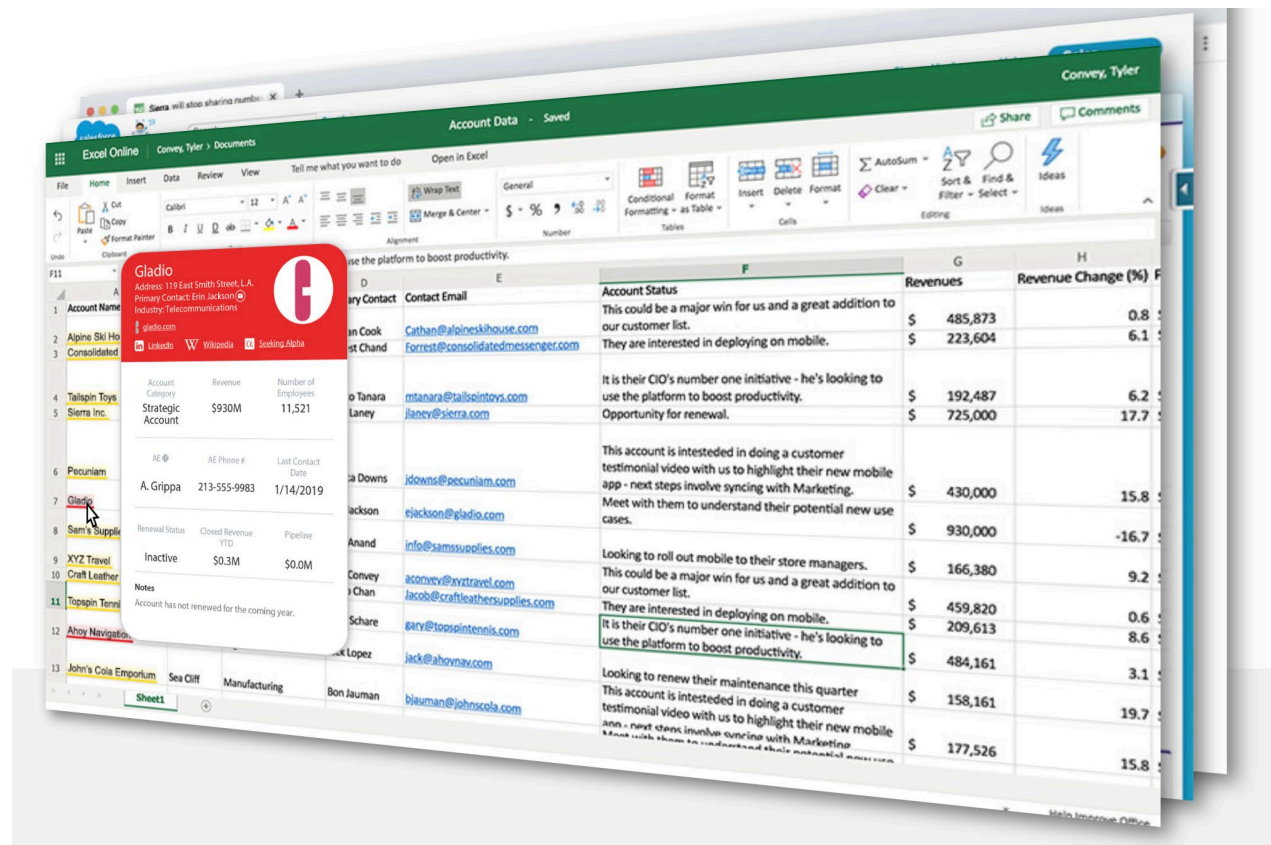
HYPERSCREEN



HYPERVISION

# HyperIntelligence - HyperCard

- Get answers by simply hovering over text—no interruptions, no delays.
- HyperCard deliver insights to users in the applications they're already using.
- Inject intelligence directly into web browsers, email clients, Salesforce.com, Slack, Office365 apps, other BI tools, and more.



# HyperIntelligence - HyperVoice

- Integrate MicroStrategy with Alexa or Google Home to turn analytics applications into a voice-enabled personal coach.
- Boost productivity by blending insights into existing workflows.

Alexa, ask MicroStrategy how my pipeline looks for this quarter?

Your pipeline for this quarter is \$30M, with your largest opportunity being \$3M with WindFarm Technologies. Would you like me to continue?



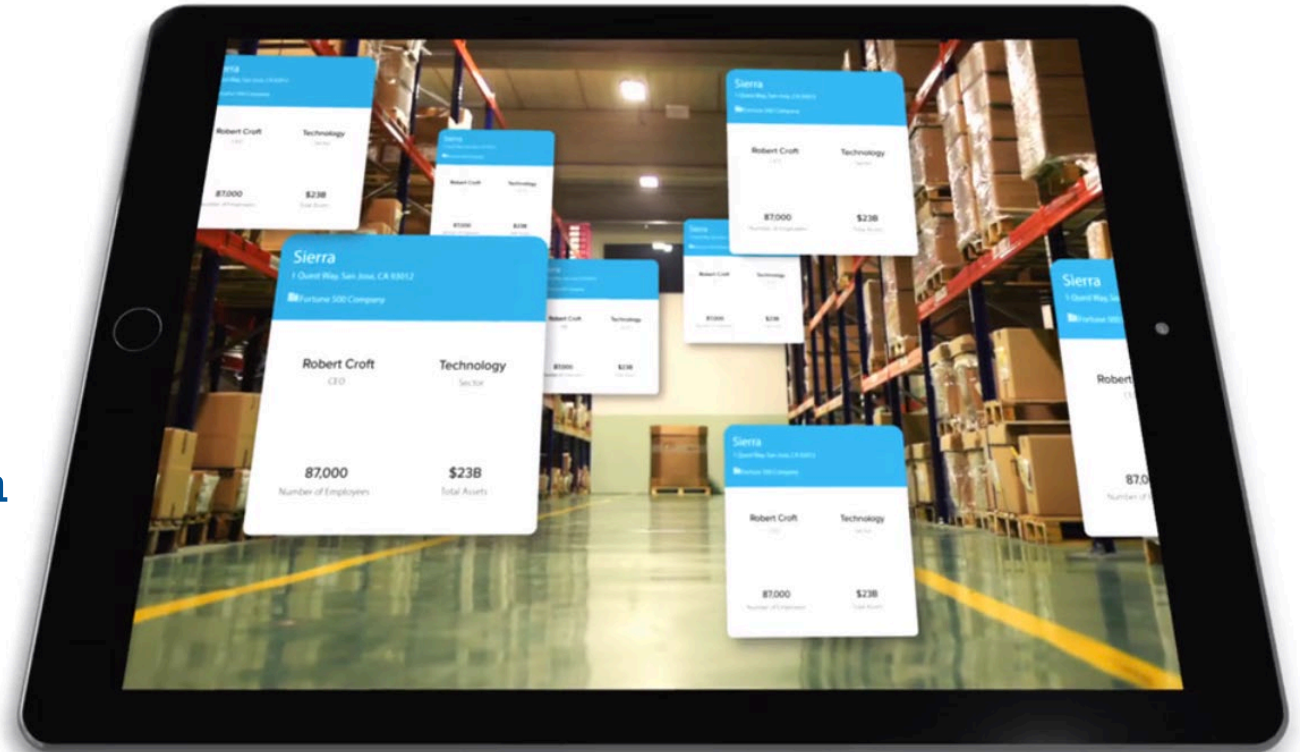
- ## DC Government





# HyperIntelligence - HyperVision

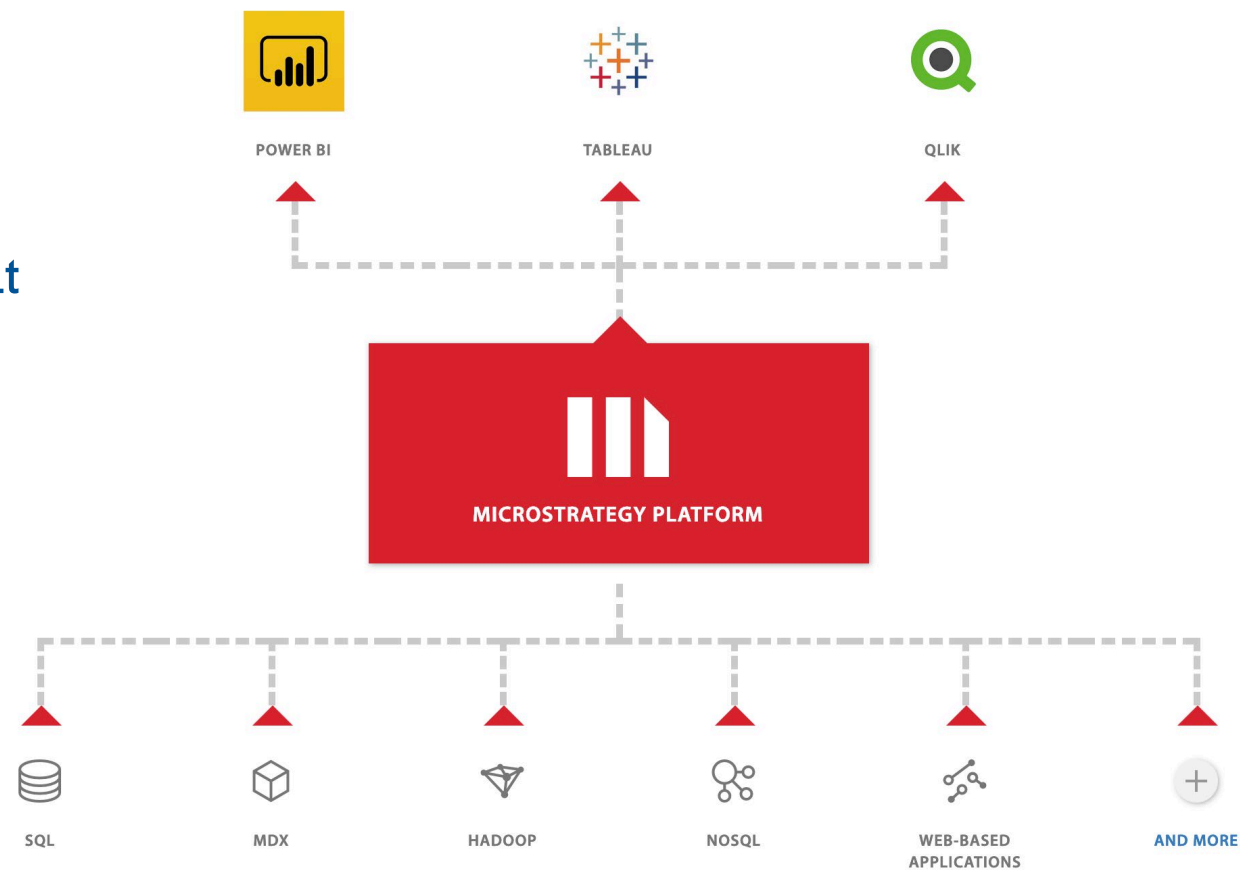
- Hold up your device for an augmented reality experience that overlays insights on objects and surroundings.
- Use image recognition to scan products and check inventory, or use GPS location to get an instant customer brief as you walk into your meeting.





# Federated Analytics

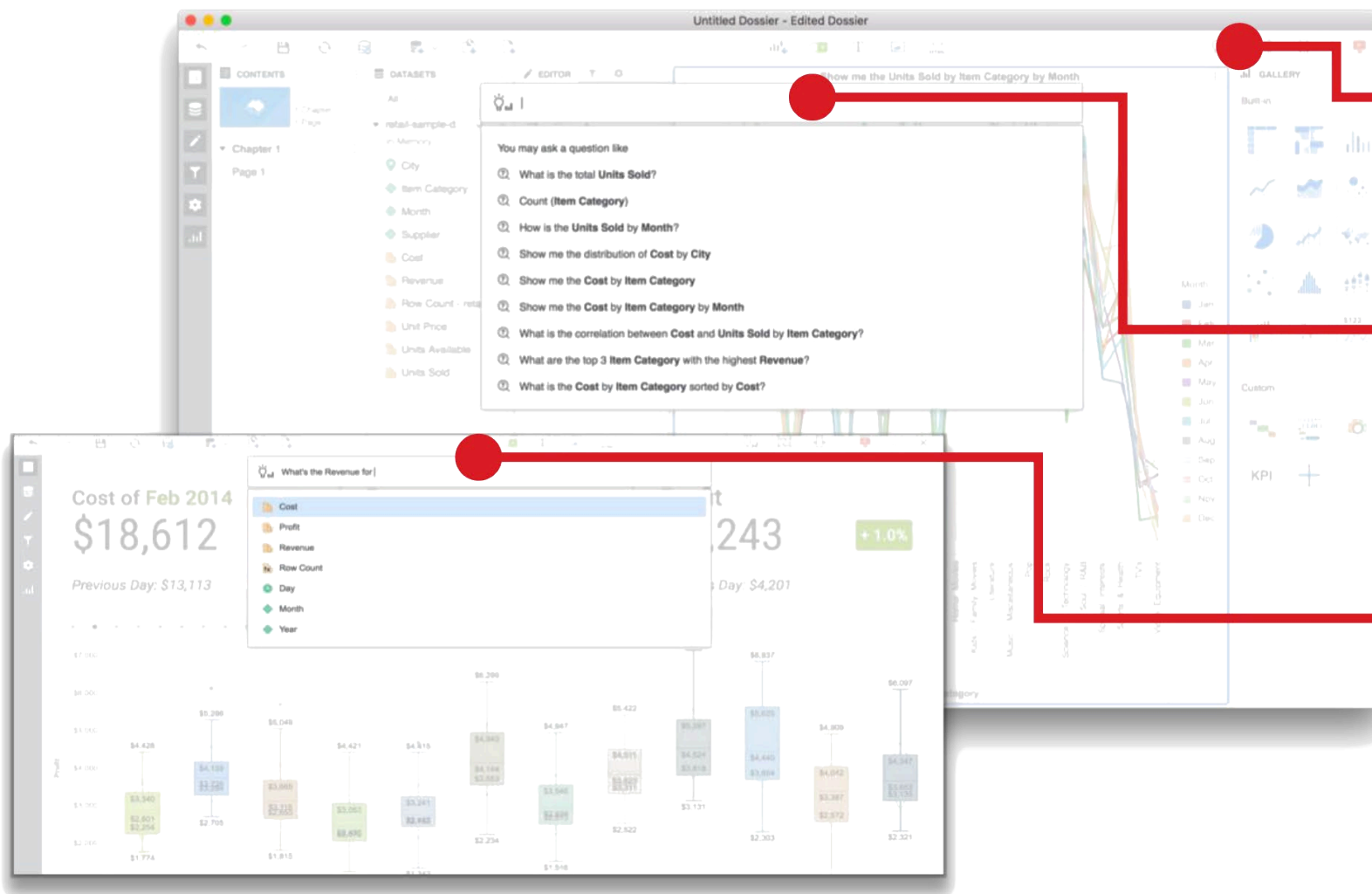
- Analysts and business teams tap into trusted enterprise data, blend in other sources, and quickly build content--without giving up their favorite self-service tools.
- Tableau, PowerBI, and Qlik users can connect directly with MicroStrategy to enhance their departmental apps with scalability, security, and data governance. It's the best of both worlds.



DC Government

# Natural Language Queries

- Natural Language Queries on google-like search interface



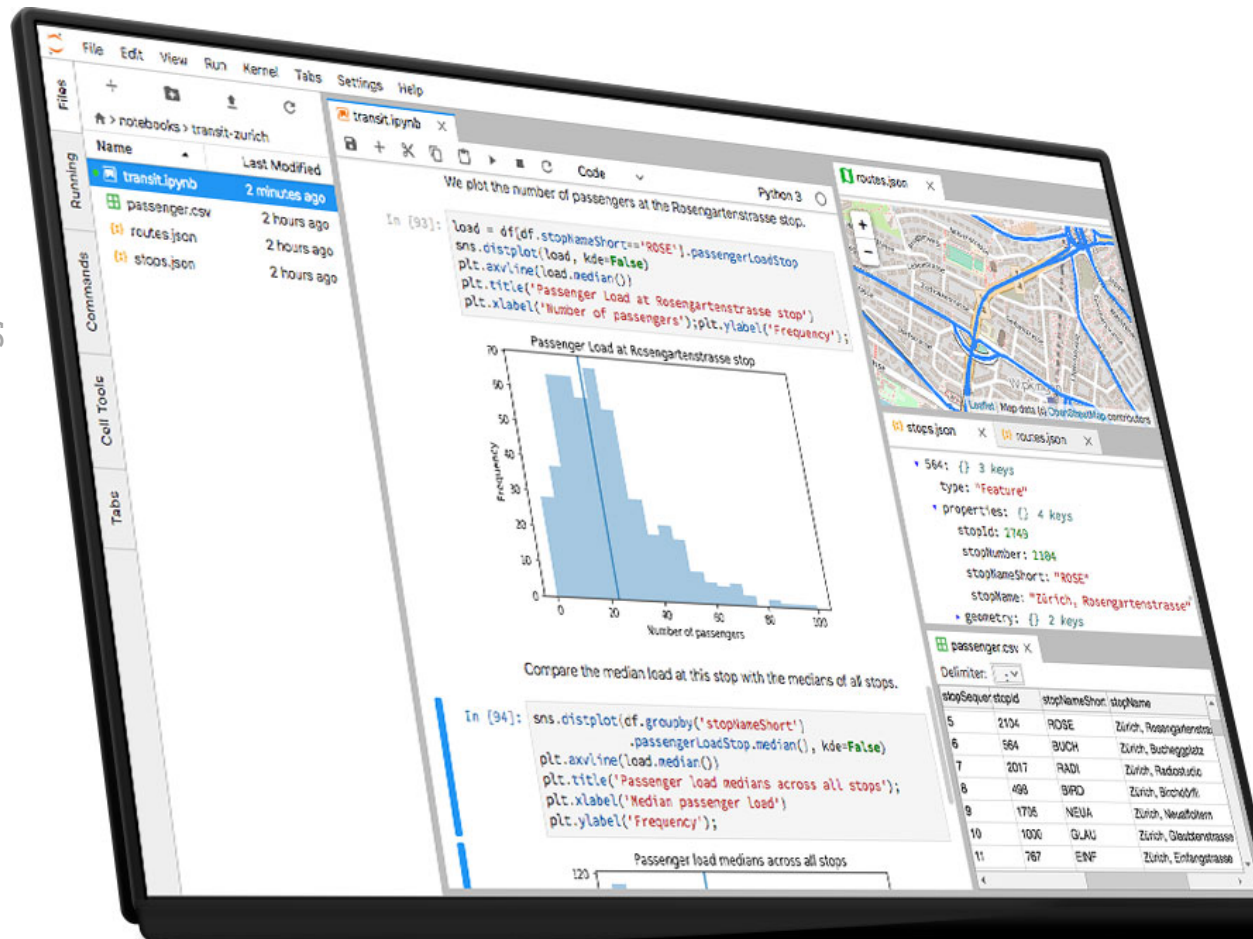
Ask questions or **type** business terms

**Recommended** questions based on the dataset

**Dynamically** suggest dataset objects as the user types a query

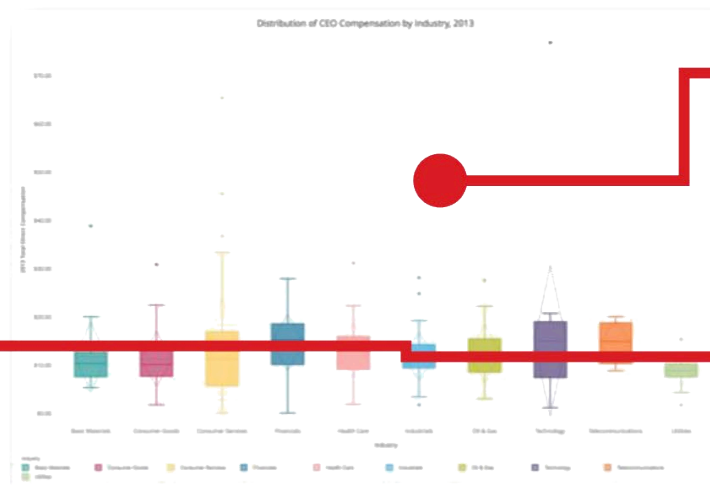
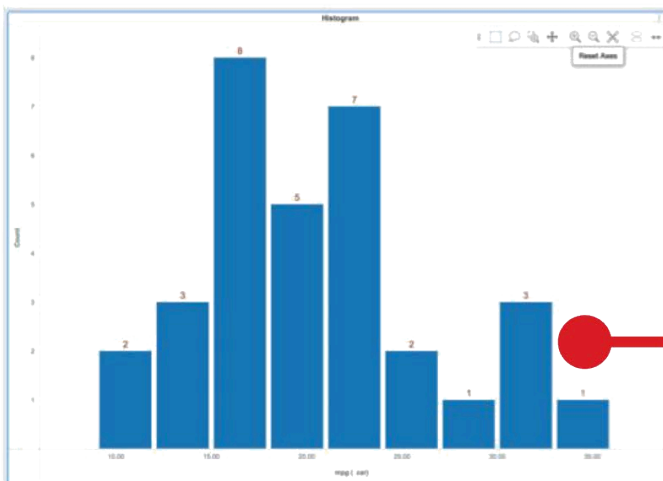
# New R and Python Packages

New integrations for R and Python are available via two open-source packages that allow data scientists to consume intelligent cubes, create R and Python algorithms, and publish results back into MicroStrategy without leaving their data science tools.



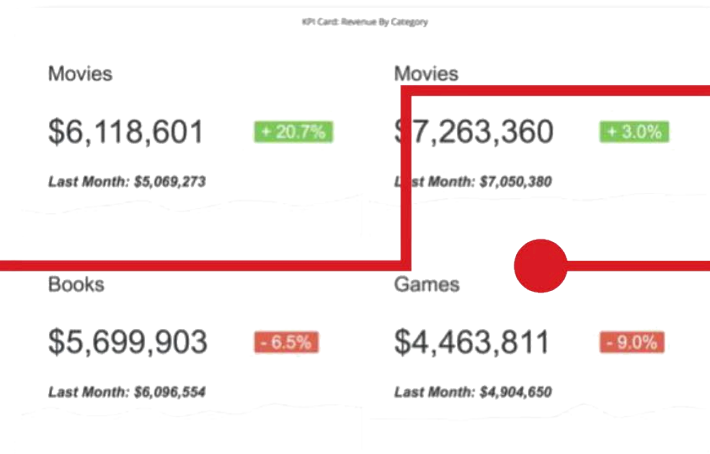
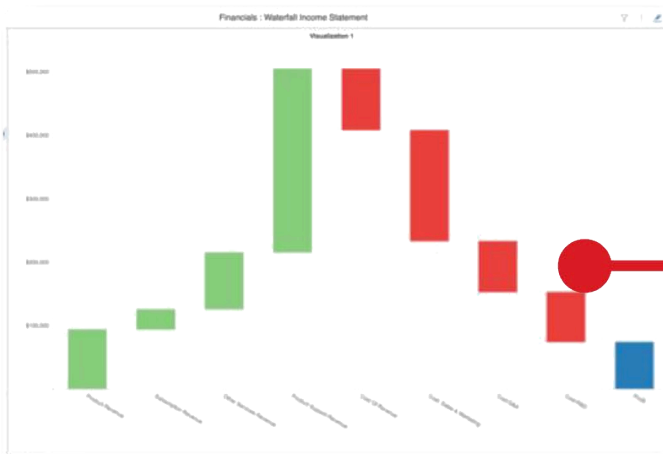
# New Out of the Box Visualizations

- Natural Language Queries on google-like search interface



**Box plot** for distribution of data based on quartile and median values

**Histogram** to reveal the underlying frequency distribution of a metric.



**Waterfall** to show the cumulative effect of sequentially introduced values

**KPI** to display the current value of any KPI. Break KPI by any attribute to see multiple cards for all elements

# Open Discussion

Next Inter-Agency Data Team  
meeting tentatively scheduled for  
end of April.

Thank you 😊

