

<b>REQUEST FOR COUNCIL ACTION</b> CITY OF SAN DIEGO				CERTIFICATE NUMBER (FOR COMPTROLLER'S USE ONLY)	
TO: CITY COUNCIL		FROM (ORIGINATING DEPARTMENT): Performance and Analytics		DATE: 9/9/2015	
SUBJECT: City of San Diego Open Data Implementation Update					
PRIMARY CONTACT (NAME, PHONE): Maksim Pecherskiy, 16195334817			SECONDARY CONTACT (NAME, PHONE): Almis Udrys, 16192365929		
<b>COMPLETE FOR ACCOUNTING PURPOSES</b>					
FUND					
FUNCTIONAL AREA					
COST CENTER					
GENERAL LEDGER ACCT					
WBS OR INTERNAL ORDER					
CAPITAL PROJECT No.					
AMOUNT	0.00	0.00	0.00	0.00	0.00
FUND					
FUNCTIONAL AREA					
COST CENTER					
GENERAL LEDGER ACCT					
WBS OR INTERNAL ORDER					
CAPITAL PROJECT No.					
AMOUNT	0.00	0.00	0.00	0.00	0.00
COST SUMMARY (IF APPLICABLE): N/A - Informational Only					
<b>ROUTING AND APPROVALS</b>					
CONTRIBUTORS/REVIEWERS:		APPROVING AUTHORITY	APPROVAL SIGNATURE	DATE SIGNED	
Liaison Office		ORIG DEPT.	Udrys, Almیس	09/10/2015	
		CFO			
		DEPUTY CHIEF	LoMedico, Stacey	09/17/2015	
		COO			
		CITY ATTORNEY			
		COUNCIL PRESIDENTS OFFICE			
PREPARATION OF:	<input type="checkbox"/> RESOLUTIONS	<input type="checkbox"/> ORDINANCE(S)	<input type="checkbox"/> AGREEMENT(S)	<input type="checkbox"/> DEED(S)	
N/A - Informational Only					
STAFF RECOMMENDATIONS: Informational Item for Committee					
SPECIAL CONDITIONS (REFER TO A.R. 3.20 FOR INFORMATION ON COMPLETING THIS SECTION)					
COUNCIL DISTRICT(S):					
COMMUNITY AREA(S):					
ENVIRONMENTAL IMPACT:					
CITY CLERK					

INSTRUCTIONS:	
---------------	--

**COUNCIL ACTION**  
**EXECUTIVE SUMMARY SHEET**  
CITY OF SAN DIEGO

DATE: 9/9/2015

ORIGINATING DEPARTMENT: Performance and Analytics

SUBJECT: City of San Diego Open Data Implementation Update

COUNCIL DISTRICT(S):

CONTACT/PHONE NUMBER: Maksim Pecherskiy/16195334817

**DESCRIPTIVE SUMMARY OF ITEM:**

Sections 2.4 and 2.5 of the City of San Diego (City) Open Data Policy require the Chief Data Officer to provide an annual written report to the Mayor and Council regarding the implementation of the Open Data Program.

This inaugural 2015 report on Open Data Program Implementation contains the findings from the City's first Data Inventory including an overview of the process, the results, and next steps. In addition, we discuss the progress the program has made in community involvement and recognition. It will be provided to the Mayor and Council annually to keep them updated on the program.

**STAFF RECOMMENDATION:**

Informational Item for Committee

**EXECUTIVE SUMMARY OF ITEM BACKGROUND:**

**Background**

The City collects, produces, and stores an enormous amount of data. A few examples include:

- \* Crime data
- \* Parking meter data
- \* Traffic counts
- \* Street paving activity

The data currently reside in a range of isolated and non-integrated data repositories, in multiple incompatible formats, and are managed by non-interoperable platforms. The data have the potential to drive deeper understanding of civic activity and policy, drive transparency and accountability, and make the City a more efficient and data-driven leader on the global civic landscape.

Properly enhanced and managed, the data has the potential to:

- \* Facilitate decision-making by staff and policymakers.
- \* Increase transparency, efficiency, and accountability.
- \* Improve information sharing and use by City employees.
- \* Allow the community to conduct analyses and build applications on top of the City's data.

The City is committed to guiding this potential towards reality.

## Open Data Program Goals

1. Increase data literacy within the City and with those who interact with the City.

There are various levels of understanding of what are data, metadata, Open Data, etc.

2. Manage data as an asset, formalizing existing oversight and ownership.

Take existing data management and responsibility structures, build on them and formalize them to create a robust data management program that would prevent data duplication and rework, as well as ensure sustainability of Open Data release.

3. Release data to the public via a strategic and predictable process.

Data releases must account for a multitude of factors including communications with stakeholders, data integrity, personally identifiable information (PII), and ensuring that we describe and release data in a way that meets Open Format specifications.

4. Ensure that data are well described and catalogued.

Understanding what data we have is crucial to both Open Data and data management. We worked with departments to create an initial inventory, and now we need to provide processes for how it can stay current.

5. Support increased use of data in decision making, as well as innovations in Open Data use.

We have a talented City workforce that makes data-driven decisions every day. In addition, fostering and expanding interest in government data from residents and other users is crucial to delivering the benefits of an Open Data program.

## Where We Are Today

In the last 6 months, we have made significant progress on all 5 of our goals:

\* Consistent with the policy timeline, we issued the inventory guidelines on March 9, before the March 31, 2015 deadline.

\* By creating the City's first ever Data Inventory, we were able to:

\* Collect base level attributes about the City's data. (Goal 4)

\* Create a base knowledge of information at the City's disposal. (Goals 2, 4, 5)

\* Establish a competent group of information coordinators across the departments (All Goals)

\* Identify individuals currently working with specific sets of data. (Goal 2)

\* Begin to establish a common body of knowledge in the City about what data are. (Goal 1 and 5).

\* Understand which of the City's software systems are authoritative for various pieces of information. (Goal 2, 4, 5)

\* Consistent with the policy, a draft version of the Technical Guidelines is under review prior to release. (All Goals)

\* We are enhancing our engagement with the civic hacking community and other interested stakeholders. For example, the City participated in the National Day of Civic Hacking by co-hosting three events on June 5th and 6th. (Goal 5)

#### Next Steps

- \* Perform a thorough analysis of inventory data.
- \* Procure an Open Data Portal.
- \* Make the inventory maintenance process routine and more efficient.
- \* Continue to provide context to City data and engage with the community.
- \* Continue refinement and updates of technical guidelines.
- \* Identify high-value datasets

#### CITY STRATEGIC PLAN GOAL(S)/OBJECTIVE(S):

Goal #1: Provide High Quality Public Service

Objective #2: Improve external and internal coordinator and communication.

Objective #4: Ensure equipment and technology are in place so that employees can achieve high quality public service.

Goal #2: Work in partnership with all our communities to achieve safe and livable neighborhoods.

Objective #5: Cultivate civic engagement and participation

Goal #3: Create and sustain a resilient, economically prosperous City.

Objective #1: Create dynamic neighborhoods that incorporate mobility, connectivity, and sustainability.

Objective #5: Enhance San Diego's global standing.

#### FISCAL CONSIDERATIONS:

N/A

#### EQUAL OPPORTUNITY CONTRACTING INFORMATION (IF APPLICABLE):

N/A

PREVIOUS COUNCIL and/or COMMITTEE ACTION (describe any changes made to the item from what was presented at committee):

On December 16, 2014, the Council passed the City's Open Data Policy with Mayor Faulconer's strong support.

#### COMMUNITY PARTICIPATION AND PUBLIC OUTREACH EFFORTS:

As part of the Open Data Program, we are working with Open San Diego, a civic-hacking focused community group. In addition the city has participated in various hackathons and tech-related events.

## KEY STAKEHOLDERS AND PROJECTED IMPACTS:

- \* Allow taxpayers to benefit from a more efficient, agile government that is capable of being responsive to them and able to deliver more with the same level of resources.
- \* Empower people to build applications that assist residents, start businesses around those applications, and contribute to the overall economic development of the City.
- \* Empower consumers of City data-based applications to have the most up-to-date, accurate and relevant information at their fingertips.
- \* Allow City employees to be more efficient and innovative by allowing them to get data from other departments more quickly and in an easily consumable format.
- \* Be recognized globally as a city government that reflects and supports the innovation of our surrounding community.

Udrys, Almis  
Originating Department

LoMedico, Stacey  
Deputy Chief/Chief Operating Officer

# City of San Diego Open Data Implementation Update



July 1, 2015



Performance  
& Analytics

# **Table of Contents**

---

1. [About This Report](#)
2. [Executive Summary](#)
3. [Inventory Process](#)
4. [Inventory Findings](#)
5. [Lessons Learned](#)
6. [Next Steps](#)
7. [Prioritization](#)
8. [Community Involvement](#)
9. [Acknowledgements](#)
10. [Attachments](#)
  - i. [Core Inventory](#)
    - i. [Step 1 - Datasources](#)
    - ii. [Step 2 - Dataset Brainstorm Guide](#)
    - iii. [Step 3 - Dataset Attribute Definitions](#)
11. [Other Accomplishments](#)
12. [Glossary](#)

# City of San Diego

---

## About this Report

---

Sections 2.4 and 2.5 of the City of San Diego (City) Open Data Policy require the Chief Data Officer to provide an annual written report to the Mayor and Council regarding the implementation of the Open Data Program. This report is available [online](#) or as a [PDF](#), and is the first City report to be written using [Git](#) and [GitBook](#), an [Open Source](#) book format and publishing toolchain.

This inaugural 2015 report on Open Data Program Implementation contains the findings from the City's first Data Inventory including an overview of the process, the results, and next steps. In addition, we discuss the progress the program has made in community involvement and recognition. It will be provided to the Mayor and Council annually to keep them updated on the program.

The inventory process was well received, with 100% of invited departments participating. In our initial run, we identified over 2,000 datasets. However, some of these datasets may not make it into the Open Data [portal](#), expected to be launched by July 1, 2016, for a variety of reasons we describe in the report including the need to process and prioritize the data for public release, which we will begin during the first half of F Y 16.

The City is well on its way towards achieving the goals of the Open Data Policy. In addition, we are on track to build a robust data management program--one that ensures the data we publish and maintain is of high quality and up-to-date.

This is not an easy effort and it would not have been possible without the vision of Mayor Faulconer and the City Council, as well as the support of the City's Chief Operating Officer Scott Chadwick, the Department of IT, and the City's management team. Through a collaborative approach, we have put together what may be the largest and most comprehensive Data Inventory effort of any government in this region. We look forward to continuing to work together on implementing the Open Data program and allowing the City to reflect and support the innovation of our surrounding community.



---

Almis Udrys,  
Director of Performance & Analytics

---

Maksim Pecherskiy,  
Chief Data Officer

## **Executive Summary**

---

The City of San Diego is committed to the principles of open, accessible, efficient, and transparent government, as well as the use of technology to help put those principles into practice. Mayor Faulconer created the Performance & Analytics Department on July 1, 2014. Its purpose is to focus on citywide efficiency, accountability, and transparency initiatives. On December 16, 2014, the Council passed the City's Open Data Policy with his strong support.

Opening up the City's data ties in directly with each of our Strategic Goals and allows us to monitor our progress in achieving them:

- Provide High Quality Public Service
- Work in partnership with all our communities to achieve safe and livable neighborhoods.
- Create and sustain a resilient, economically prosperous City.

For additional background and history of the Open Data Program in the City, please refer to the following:

<http://bit.ly/sdodpres>.

The guiding principle for the program is to get the proper stakeholders with the right skills, involved in a timely manner, equipped with the appropriate technology and accurate data to facilitate good decisions and innovative solutions for our residents.

## **Background**

---

The City collects, produces, and stores an enormous amount of data. A few examples include:

- Crime data
- Parking meter data
- Traffic counts
- Street paving activity

The data currently reside in a range of isolated and non-integrated data repositories, in multiple incompatible formats, and are managed by non-interoperable platforms. The data have the potential to drive deeper understanding of civic activity and policy, drive transparency and accountability, and make the City a more efficient and data-driven leader on the global civic landscape.

Properly enhanced and managed, the data has the potential to:

- Facilitate decision-making by staff and policymakers.
- Increase transparency, efficiency, and accountability.
- Improve information sharing and use by City employees.
- Allow the community to conduct analyses and build applications on top of the City's data.

The City is committed to guiding this potential towards reality.

## Benefits of Open Data

---

1. Allow taxpayers to benefit from a more **efficient**, agile government that is capable of being responsive to them and able to deliver more with the same level of resources.
2. **Empower** people to build applications that assist residents, start businesses around those applications, and contribute to the overall **economic development** of the City.
3. **Empower** consumers of City data-based applications to have the most up-to-date, accurate and relevant information at their fingertips.
4. Allow City employees to be more **efficient** and innovative by allowing them to get data from other departments more quickly and in an easily consumable format.
5. Be recognized globally as a city government that reflects and supports the innovation of our surrounding community.

## Open Data Program Goals

---

### 1. Increase data literacy within the City and with those who interact with the City.

There are various levels of understanding of what are data, [metadata](#), Open Data, etc.

### 2. Manage data as an asset, formalizing existing oversight and ownership.

Take existing data management and responsibility structures, build on them and formalize them to create a robust data management program that would prevent data duplication and rework, as well as ensure sustainability of Open Data release.

### 3. Release data to the public via a strategic and predictable process.

Data releases must account for a multitude of factors including communications with stakeholders, data integrity, personally identifiable information (PII), and ensuring that we describe and release data in a way that meets [Open Format](#) specifications.

### 4. Ensure that data are well described and catalogued.

Understanding what data we have is crucial to both Open Data and data management. We worked with departments to create an initial inventory, and now we need to provide processes for how it can stay current.

### 5. Support increased use of data in decision making, as well as innovations in Open Data use.

We have a talented City workforce that makes data-driven decisions every day. In addition, fostering and expanding interest in government data from residents and other users is crucial to delivering the benefits of an Open Data program.

## Where we are today

---

In the last 6 months, we have made significant progress on all 5 of our goals:

- Consistent with the policy timeline, we issued the inventory guidelines on March 9, before the March 31, 2015 deadline.
- By creating the City's first ever Data Inventory, we were able to:
  - Collect base level attributes about the City's data. (Goal 4)
  - Create a base knowledge of information at the City's disposal. (Goals 2, 4, 5)
  - Establish a competent group of information coordinators across the departments (All Goals)
  - Identify individuals currently working with specific sets of data. (Goal 2)
  - Begin to establish a common body of knowledge in the City about what data are. (Goal 1 and 5).
  - Understand which of the City's software systems are authoritative for various pieces of information. (Goal 2, 4, 5)
- Consistent with the policy, a draft version of the Technical Guidelines is under review prior to release. (All Goals)
- We are enhancing our engagement with the civic hacking community and other interested stakeholders. For example, the City participated in the [National Day of Civic Hacking](#) by co-hosting three events on June 5th and 6th. (Goal 5)

## Next Steps

---

- Perform a thorough analysis of inventory data.
- Procure an Open Data [Portal](#)
- Make the inventory maintenance process routine and more efficient.
- Continue to provide context to City data and engage with the community.
- Continue refinement and updates of technical guidelines.
- Identify high-value datasets and prioritize data for release.
- Review and potentially update Open Data Policy.
- Begin to release data by July 1, 2016.

# Inventory Process

---

## Purpose

---

The purpose of the Data Inventory is to get a bird's eye view of data throughout our City departments that will assist stakeholders in identifying datasets they consider high value for publication. However, we deliberately designed the inventory process to do much more, as follows:

- Establish a base level inventory of the data the City has at its disposal.
- Create and collect a base level of attributes describing our City's data.
- Establish a competent group of information coordinators across the departments.
- Identify people currently working with and managing data.
- Understand which of the City's software systems are authoritative for various pieces of information.
- Identify critical data and be able to prioritize it for release or internal access.
- Understand which of the City's software systems are authoritative for various pieces of information.

## Process Overview

---

The first part of the inventory was to identify information coordinators from each of the 35 departments/programs we engaged. Most departments assigned a single point of contact to work with the Chief Data Officer. However, several larger departments chose to assign multiple coordinators to represent certain divisions. We sought out information coordinators that are knowledgeable in the business level--of how the department operates, and how users make decisions based on that knowledge. In total, we worked with 65 coordinators.

We designed the inventory process mindful of the following realities:

- The information coordinators' time must be used as efficiently as possible, as this process is not their primary responsibility.
- Coordinators may not have a high level of technical knowledge.
- Execution time frame is only three months.
- Coordinators do not have the same base level of understanding of what is "data".

Consequently, we emphasized the following:

- Minimize large in-person meetings, but be available for small face-to-face meetings, maintaining preference for online collaboration.
- Provide all the necessary information to everyone all the time.
- Continuously collect feedback, and be flexible enough to adjust the process.
- Automate repeatable tasks and allow technology to increase execution velocity.
- Technology or the CDO (Chief Data Officer) must not be a bottleneck to any of the coordinators.
- Process must be as efficient as possible to minimize rework and increase accuracy.
- Break down large chunks of work into smaller ones. Operate in short spurts, but allow for full-scope completion all at once.

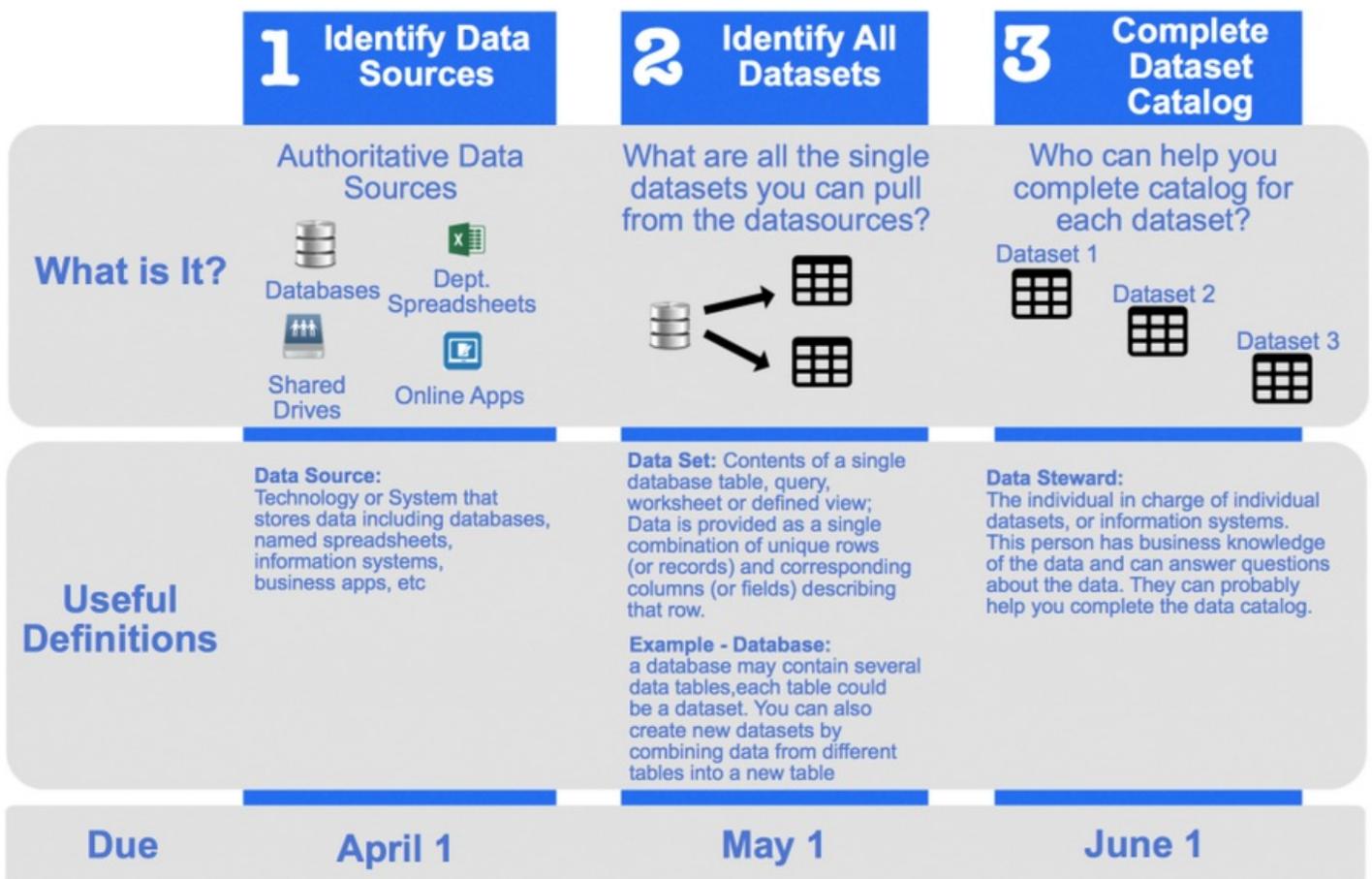
- Create infrastructure for future growth and maintenance.
- Track progress.

## Kickoff

Performance & Analytics conducted two kickoff meetings with the same content to field preliminary questions, engage as many coordinators as possible, and introduce the process. These 3-hour meetings were the only large meetings held during the inventory process.

## Inventory Process

We broke the inventory down into three steps, giving coordinators one month to complete each step. We gave each department a spreadsheet to fill out with all steps and explanations built into the sheet.



## Step 1

For Step 1, we asked coordinators to identify any authoritative datasources their department uses. We identified a datasource as any system that stores the primary version of data. For example:

- Information Systems (such as public facing applications sandiego.gov or internally used applications like SAP)
- Databases (such as Microsoft Access and Business Warehouse)
- Excel Spreadsheets on Shared Drives or Personal Desktops
- Access Databases on Shared Drives or Personal Desktops
- 3rd Party Vendors and data hosted on vendor systems (such as Survey Monkey or Google Analytics).

For a further description of what we provided as datasource definitions, please see [Core Inventory, Step 1 - Datasources in the Attachments Section](#).

In addition to identifying datasources, we asked coordinators to identify:

- The person that is responsible for the technical management of the system (IT Subject Matter Expert).
- The person that has the business level knowledge and understanding of what data are stored within the system (Business Subject Matter Expert).
- Any potential technical changes (replacements, upgrades) the system might experience in the future.

## Step 2

In Step 2, coordinators identified potential datasets within the datasources from Step 1. We designed this exercise to explore what kind of data are stored within the datasources. Because each department has its own culture and style of communication, we provided a brainstorm guide to allow coordinators to come up with datasets in as many ways as possible. The brainstorm guide can be found in [Core Inventory, Step 2 - Dataset Brainstorm Guide in the Attachments Section](#).

## Step 3

In Step 3, we collected information needed to get a basic understanding of datasets, in order to ultimately assist the City in prioritizing them for release. We asked about the following indicators:

- Brief description of data
- Frequency of data change
- Format
- Existing publication
- Link to existing publication
- Priority/value
- Data Classification (Public / Protected / Sensitive)
- Data Quality Concerns
- PRA frequency
- Comments / Extra Info

Please review the [Core Inventory, Step 3 - Dataset Attribute Definitions in the Attachments Section](#) for the definitions of the fields above.

## Technology

### Inventory Sheets

Working closely with the Department of Information Technology, we took advantage of technology to complete the inventory in the required timeframe.

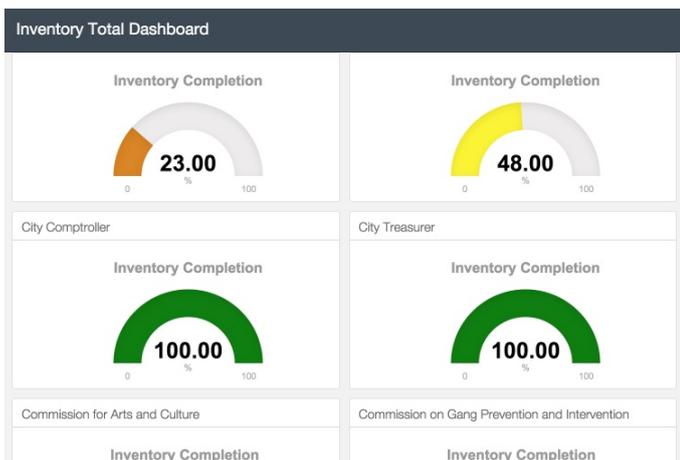
Through the use of technology in the inventory process, we were able to facilitate the following:

- Automate creation of spreadsheets per department.
- Maintain centralized access to all sheets while giving coordinators access to their own.
- Increase our ability to adjust quickly to feedback by writing code that made changes to all the department spreadsheets at once.
- Allow coordinators to edit their inventory documents simultaneously in real time, which prevents wasting time on merging changes between document versions.
- Aggregate and track progress for all departments in one location.

### Reporting

Department	Normalized Ratio	Notify Step 3 = Complete	Step 3 Manual Multiplier	Step 1 %	Step 2 %	Step 3 %	Inventory %
Communications	1	1	100	100.00	100.00	100.00	100.00
Corporate Partnerships and Development	1	1	100	100.00	100.00	100.00	100.00
Debt Management	1	1	100	100.00	100.00	100.00	100.00
Development Services	1	1	100	100.00	100.00	100.00	100.00
Docket	1	1	100	100.00	100.00	100.00	100.00
DoIT	1	1	100	100.00	100.00	100.00	100.00
Economic Development	1	1	100	100.00	100.00	100.00	100.00
Environmental Services	1	1	100	100.00	100.00	100.00	100.00
Ethics Commission	1	1	100	100.00	100.00	100.00	100.00
Financial Management	1	1	100	100.00	100.00	100.00	100.00
Fire-Rescue Department	1	1	100	100.00	100.00	100.00	100.00

To facilitate continuous reporting and track progress for all departments simultaneously, we built a central dashboard.



However, department coordinators did not need the same level of depth, so we provided them a simplified interface.

### Dissemination of information

One of our main goals was to make online resources available to all coordinators all the time. For disseminating the information, we chose to use a technology called [GitBook](#), which is an [open source](#) publishing work-flow system. It also allowed us to continuously incorporate feedback and update the instructions without having to redistribute them to coordinators.

In addition to the benefit of receiving continuous feedback from City staff, we also analyzed how the staff used the documents. By continuously tracking our documentation's effectiveness, we issued targeted clarifications and improved internal communication around Open Data. For example, earlier versions of the inventory had separate examples for each step. We noticed that many coordinators went to see the examples, but left those pages quickly, which suggests that they did not find them useful. When we changed to a single example that covered the entire inventory, we instantly observed people staying on pages longer and getting more use from the examples.

**Monitoring the navigation flow of users:**

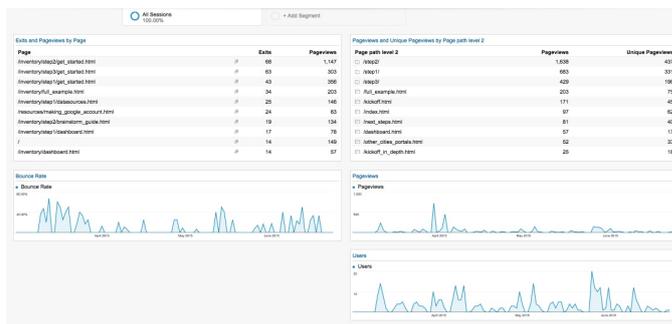


We saw how people navigated through the inventory: which pages led them to other pages. Based on the feedback, we made several decisions on how to lay out the guidelines.

**Tracking various metrics to gauge user behavior and adjust accordingly:**

We tracked:

- Pages where the most people exited the guidelines.
- Viewed pages of the guidelines.
- Ratio of how many people left the site after viewing only one page.
- Pageviews, users and returning users.



**Tracking pageviews per session and returning users**



In order to gauge the usefulness of the guidelines, it was important to know how many people spent a significant amount of time reading the guidelines, and how many people returned.

## Compiling

R Script handles inventory data compilation and reporting across the department spreadsheets. R Script pulls summary inventory data from all spreadsheets, aggregates and then cleans the data. In order for the inventory to be effective, it must be continuously updated, and these tools allow us to do so now and in the future.

## Additional Sources

In addition to the manually generated inventory, we are working with the Department of IT to take advantage of available technological data collection methods. We will rely on the following sources to verify and augment the inventory data over the coming year:

- List of [GIS](#) (Geographic Information Systems) Data
- List of Business Warehouse Reports
- List of Data provided by CGI (the application support sourcing vendor)
- Automated scanning of sandiego.gov for potential datasets

# Inventory Findings

The Data Inventory we compiled is the first Data Inventory in the history of the City. It is a living document and will continue to evolve. Since this first ever inventory is manually generated, it will require cleaning, as there are a number of values that do not fit into the validation criteria (for example: a coordinator inputting "publicly available" instead of "public" in the data classification field). In addition, we still need to validate and merge information from non-human sources. Finally, before release, we must mitigate data that may expose PII or cybersecurity vulnerabilities. After we have verified the inventory data, we will post a list of datasets for public review.

## Submissions By Department

Below is a list of submissions by department. Departments vary in size, number of employees, budget, and nature of workload, so we expected a wide diversity in these figures. They do not indicate the quality of any department's participation in the process. As we proceed with the program, the inventory will become more robust and these numbers will grow more accurate.

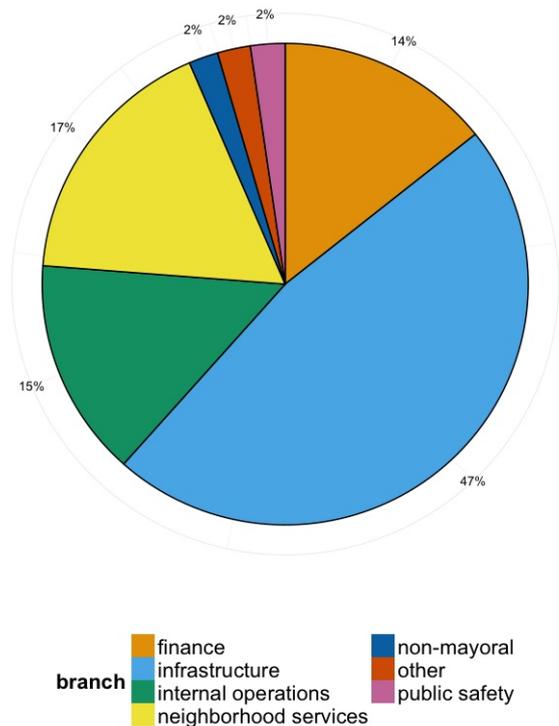
### Submissions Per Department

<b>dept</b>	<b>numDatasets</b>
ada compliance and accessibility (office of)	8
citizens' review board on police practices	9
city auditor	6
city clerk	30
city comptroller	93
city treasurer	63
commission for arts and culture	27
commission on gang prevention and intervention	3
communications	4
corporate partnerships and development	6
debt management	64
development services	24
docket	2
doit	247
economic development	11
environmental services	112
ethics commission	1
financial management	31

fire-rescue department	24
homeland security	11
human relations commission	6
human resources	22
library	38
park and recreation	221
performance & analytics	8
personnel	3
planning department	16
police	12
public utilities	749
public works / ecp	8
purchasing & contracting	17
read	14
risk management	42
special events	17
transportation & storm water	99
<b>Total</b>	<b>2048</b>

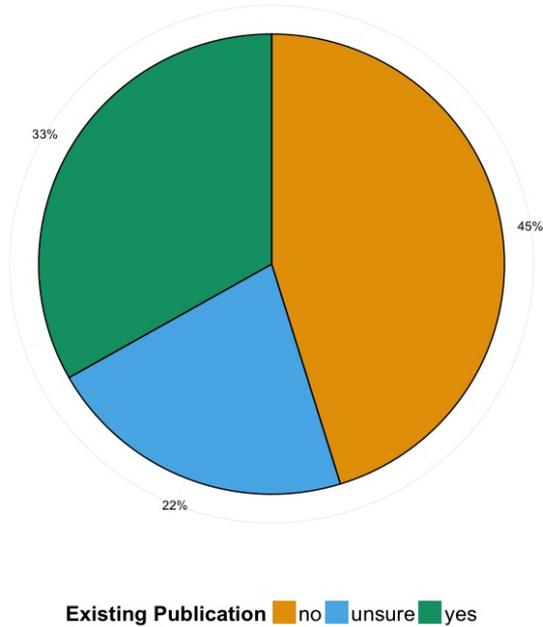
### Submissions By Branch

The City has a substantial amount of data, and information coordinators worked hard to educate Performance and Analytics about the data.



## Public Dataset Analysis

As part of the inventory process, we asked departments to identify datasets they consider to be already public, and if they are already publishing them. Because this information is preliminary and still needs to be validated, below is a general citywide aggregated overview of department-reported information about public datasets.



### Citywide Publication of Reported Public Datasets

Departments already consider nearly half of the reported datasets to be public but are not yet publishing them. There are others that they do publish, but not in a central, organized location and in an [open format](#). As we prioritize public datasets for release, we envision these to be among the low-hanging fruit to consider for publication in the Open Data [portal](#). Some examples include:

- Capital Improvement Projects (PW).
- Library Locations and Hours (Library).
- Event Calendars (Parks & Special Events).
- Parking Meter Data (Treasurer).

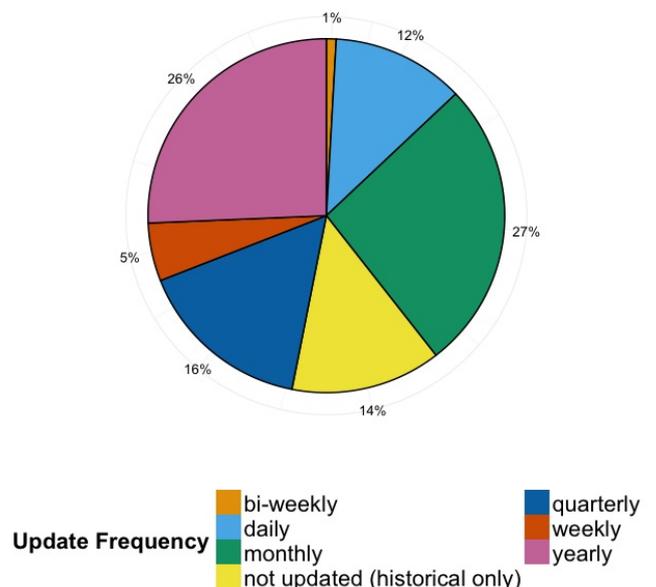
## Update Frequency

As a result of the inventory process, we also obtained a general overview of how frequently departments update datasets. Although this information has yet to be verified (and thus is presented in aggregated form below), dataset update frequency offers another criteria we will consider when identifying public datasets for release. Uploading datasets to the [portal](#) that are frequently updated would require automated processes that may require additional resources.

### Frequency of Updates of Reported Public Datasets

Public datasets updated less frequently represent other low-hanging fruit, because they may not warrant a need for an automated process to publish data. Examples of datasets already published, with low update frequency:

- Fire Hydrants (PUD)
- Storm Water Runoff Report (TSW)
- Business Improvement Districts (Economic Development)
- Rec Spaces (Park and Rec)
- Fire Stations (Fire Department)
- Annual Permit Issuances (DSD)



## **Lessons Learned**

---

- We have a plethora of amazing talent in the City. City staff are passionate about their duties, committed to customer service, and knowledgeable about the ins and outs of their jobs.
- Communication is important. We need to make sure that the way in which we communicate works for everyone, that everyone uses our tools and understands them. Analytics allow us to facilitate those results.
- City departments will save time and avoid duplication of effort when we set up a system with a continuous flow of data from other departments.
- The City currently leverages multiple technologies for maintaining and controlling data that we are looking forward to leveraging.
- The City is beginning new projects that will encompass data integration and multiple IT changes. It will be important for us to make sure that we're integrating with current and ongoing initiatives.

### **Some Feedback Received from Information Coordinators**

- "My department should select metrics more targeted towards measuring our performance."
- "My department needs to be more explicit in our data collection methodologies."
- "My department needs to do more work that targets effects we can measure. "
- "From this exercise, I believe that my department needs to actively evaluate our priorities and align them with our mission."
- "We need to increase capacity for managing and analyzing data within my department."
- "My department needs to focus on setting an end goal, and measuring our work by data to back it up."

## **Next Steps**

---

### **Perform a thorough analysis of inventory data**

Since the inventory is manually generated, we need considerable cleaning of the data in order to correct duplicative and inconsistent data entry. In addition, we have yet to merge information from additional sources to validate, and possibly supplement, today's inventory. Finally, there are risks that our data may expose PII or cybersecurity vulnerabilities. After we have verified the inventory data and taken steps to mitigate these potential vulnerabilities, we will post it for public review and to assist in the identification of high-value datasets for prioritization and release.

### **Procure an Open Data Portal**

Since it is difficult for many people to understand data when they are provided in a purely machine readable format, consistent with the Open Data Policy, in F Y 16 we will procure a software solution for making user-friendly visualizations of such information. In addition, this software will host datasets we make available to the public.

### **Make the inventory process routine and more efficient**

The initial inventory largely relied on manual input, but as we continue along the path of treating our data as valued assets, we will explore and implement options to automate inventory maintenance and further educate information coordinators.

### **Continue to provide context to City data and engage with the community**

We will continue our community outreach and engagement efforts with [Open San Diego](#). We will explore new avenues for contextualizing City data, so that they can be consumed by all interested parties. We will pursue partnerships with nonprofits or educational institutions to analyze data for the betterment of our communities and our City operations.

### **Continue refinement and updates of technical guidelines**

We designed technical guidelines to operate as a living document, and we will continue to adjust as we move along the process of publishing data and managing data. A well-designed release process--one that ensures any data released to the [portal](#) continues to be up-to-date--will be a key component of the technical guidelines, and a well-defined internal data management structure will be crucial in supporting it.

## **Identify high-value datasets and prioritize data for release**

After the inventory data are cleaned and analyzed, we will prioritize data based on the plan outlined in the technical guidelines for release. While it may be subject to revision, it is included in this report in the [next section](#). Furthermore, as we leverage Open Data for creation and future upgrading of the citywide performance dashboard, datasets supporting those indicators will be prioritized higher for release.

## **Review and potentially update Open Data Policy**

Since the Policy calls for a yearly review of itself, we will be reviewing the Open Data Policy before the end of the calendar year.

## **Begin to release data by July 1, 2016**

We are still on schedule to begin data releases by July 1, 2016, as well as looking forward to bringing in currently released data under the governance.

# Prioritization

The Open Data Policy calls for us to release high-value datasets to the public. High-value is a subjective term, highly relative to the standpoint of the person making the judgment. Therefore, to have a solid determination of high-value, we need to understand who it is that may be making the judgment:

## Components of Prioritization

### Personas

In our thinking about Open Data, we identified several high-level personas that would naturally care about what is released. These personas will be part of our evaluation on what constitutes datasets that are scheduled for release.



In order to prioritize, we plan to measure a dataset on the following overarching factors:

## Value

The value of a dataset is gauged by demand from various stakeholders; whether this dataset will increase transparency, accountability or internal efficiency; or create economic opportunity.

## Security

The City will place the highest priority on protecting data that exposes the City to security risks or the public's private information from release.

## Quality

Low-quality data (i.e. missing fields, erroneous entry, manually updated) could receive a lower prioritization, since it may need some extra work.

## Readiness

Readiness gauges the amount of work required to convert the data to an [open format](#), and whether data are already routinely published.

## Prioritization Matrix

Based on the preceding components of prioritization, we have derived an initial prioritization matrix that will be revised as necessary following a legal and technical review:

Field Name	Field Description	Prioritization Category
Dataset	Dataset Identifier	Meta
Mayor Demand / Council Demand	Is there demand from the Mayor / Council for this dataset?	Value
Interdepartmental demand	Can releasing these data positively influence workflows / performance across City silos?	Value
Departmental demand	Does the department desire that these data be released?	Value
Dataset included in Open Data Census	The Open Data Census contains some of the more highly requested datasets across the country and is a good indicator for demand ( <a href="http://sdgo.io/1IzdOom">http://sdgo.io/1IzdOom</a> )	Value
There is an application built ready to use these data.	If there is an application built on sample data that we can deliver to residents or other users quickly and provide impact.	Value
There is a known, constant stream of PRAs for these data	If releasing these data will alleviate some PRA work from departments.	Value

Resident Demand	Percentage of surveyed residents who want these data released	Value
# Of Defined Personas Affected Positively	Based on our persona definitions, which of the personas are likely to use these data? ( <a href="http://sdgo.io/1HeIOiK">http://sdgo.io/1HeIOiK</a> )	Value
Is this data already being published?	Are these data already being published, but not in a central, organized location and in <a href="#">open format</a> ?	Readiness
Coordinator Value Assignment	Did the Coordinator mark these data as high / med or low priority	Value
Data Sensitivity Assignment	Are these data public, protected or sensitive	Security
Data Quality Concerns	Are there concerns about data quality?	Quality
Data governance structure	Does a minimum viable level of data governance structure exist for this dataset?	Quality
Data frequency of change	If the data get updated often, releasing it without ETL will render it irrelevant, and building ETLs may require investment of time and money.	Quality
ETL Required?	A 3-level indicator of whether ETL is required - 1=yes, 2=yes,but can delay, 3 = no	Quality
Do this data contain potential PII (Personally Identifiable Information), or PCI (Payment Card Industry) information?	If these data contains PII or PCI information, they will need special handling making it harder to release	Security
Do these data contain information detrimental to the City's security if released or information that is business sensitive?	If these data contain such information, they will need special handling making it harder to release	Security
Do these data contain information that is public but only under specific terms?	If these data contain such information, they will need special handling making it harder to release	Security
Data extraction complexity	If it's hard to extract data from a given system, it may result in a lower prioritization	Readiness
<a href="#">Metadata</a> Availability	<a href="#">Metadata</a> Availability	Readiness
Do these data support a performance indicator?	If these data support a published performance indicator for the department, that will cause a higher level of prioritization	Value

# Community Involvement

---

## Purpose

Community outreach and engagement is critical for any Open Data effort to be successful. The local Code for America brigade, [Open San Diego](#), largely drives engagement with the community purely as a volunteer group. It is important for us to work closely with them as they are an active amplifier of our efforts.

## Open San Diego

[Open San Diego](#) was a staunch advocate for adopting the Open Data Policy and hiring a Chief Data Officer. In the past six months, [Open San Diego](#) under the leadership of Xavier Leonard, has accelerated its activities. It is now a group that we partner with on a regular basis. It has several [ongoing projects](#), such as:

- A code enforcement notification application, leveraging City data,
- A street sweeping notification application, leveraging City data,
- A parks finder application, leveraging City data.

They work with other interested stakeholders to increase membership and civic engagement with data.

In order to build proper context between the community and the City, the following City employees have presented at [Open San Diego](#), upon our request:

### **Jonathan Behnke (CIO)**

Jonathan had an informative presentation and discussion with the group about the City's IT infrastructure, forward plans and roadmap.

### **Scott Daeschner (GIS Manager)**

Scott did a commendable job describing the ecosystem of [GIS](#) (mapping) information within the City, as well as the Region.

### **Maksim Pecherskiy (CDO)**

Maksim discussed how to use [Git](#) and Github, as well as about his experience working with the local brigade in Puerto Rico.

### **DeeDee Alari, Jonathan Carey, Mary Davis (City Treasurer)**

DeeDee, Jonathan and Mary talked to the group about the City's Smart Meter program, and what data the meters collect and generate. It proved to be a lively discussion about parking, meters, their purpose and placement.

### **Almis Udrys (Director, Performance & Analytics)**

Almis has appeared before [Open San Diego](#) in the years leading up to the adoption of the Open Data Policy to provide updates on the City's efforts in this space.



These appearances enhance the relationship between the City and the civic hacking community. In addition, other City employees have become regular attendees of [Open San Diego](#) Meetings and are actively working on civic hacking projects.

### National Day of Civic Hacking

[Open San Diego](#) coordinated 3 simultaneous events on June 5 and 6, 2015 for the [National Day of Civic Hacking](#), in which many software developers and other community members got involved.



Team members from 18F, a top notch team of designers, developers and product specialists embedded in the federal General Services Administration, had representatives at the event. The City was [recognized on a national stage](#).

### CyberTech

On May 30 and 31, CyberTech, a global cybersecurity and Internet of Things network ecosystem, in partnership with the City and AT&T, hosted a [hackathon](#) to build applications of the City's already Open Data.



## Outstanding Departments

---

This is not an easy process and different from what most City staff do day-to-day. All the departments did an amazing job on the inventory, but there are some departments (and coordinators) that deserve special recognition.

They provided constructive feedback and exceptional assistance, helping to shape the inventory process. This is not based on the number of datasets each submitted, but rather on their active involvement in the process.

In no particular order:

### **Department of IT**

- Jonathan Behnke
- Gary Hayslip
- Scott Daeschner
- Margo Sanchez
- Debra Bond
- Noel Gilchrist

### **City Clerk**

- Elizabeth Maland
- Sheron Summers

### **City Treasurer**

- Mary Davis
- Dee Dee Alari
- Jonathan Carey

### **Park and Recreation**

- Andrew Field

### **Commission for Arts and Culture**

- Dana Springs
- Christine Jones

### **Financial Management**

- Ben Battaglia

### **Fire-Rescue Department**

- Ken Barnes

### **Public Utilities**

- Linda Schmidt and all the PUD coordinators.

### **Transportation and Storm Water**

- Kris McFadden
- Hassan Yousef
- Linda Marabian

## **Additional Acknowledgements**

**Joy Bonaguro (CDO of San Francisco)** Joy is a pleasure to collaborate with and was instrumental in helping plan the basics of the inventory process.

**Allen Gunn** The advice we have received from Allen (Director of Aspiriontech) has been invaluable and has led us down the path of better communication with the information coordinators.

# Step 1 - Datasources

---

## What is a Datasource?

A Datasource is a technology or system that stores data. Some examples include:

- Information Systems
- Databases
- Excel Spreadsheets on Shared Drives or Personal Desktops.
- Access Databases on Shared Drives or Personal Desktops.
- 3rd Party Vendors and data hosted on vendor systems.

## Helper Questions

In order to help you identify some datasources, use this list of helper questions:

- What databases does your department use?
- What information systems does your department use?
- What applications capture information or are used in your business processes?
- Are there some datasources kept in spreadsheets on your desktop?
- Are there some data that you work on with other people stored on shared drives?
- Are you already publishing information out on the web or in reports? Where does that information come from?
- Do you use Excel Spreadsheets or Access Databases to hold any information?

## Do I have to list every information system my department uses?

While ideally we would prefer to know every system you use that stores data, we recognize it may be overwhelming and burdensome for some departments to list every single piece of software they use that stores data. In those situations, we are asking you to use your judgment: if this is something that gets PRA'd or something that holds information you think may be useful to someone in another department please put that system in as a ***datasource***.

## Do I have to list every Spreadsheet / Access DB?

It's important to understand that we are **not** looking for a listing of every spreadsheet / Access DB you, your department or your colleagues own or use.

The spreadsheets we would like to ask you to add as Datasources are the type that are:

- Periodically Updated (Weekly / Monthly / Quarterly / Annually / etc.).
- Used to run reports or are reports.
- Used to update leadership.

This is not an exhaustive list of attributes that would make a spreadsheet a datasource, but we ask you to use your best judgment based on the examples below.

If you are hesitant about whether to include a spreadsheet / Access DB in the list, feel free to [email Maksim](#), or err on the side of caution and include it (**unless doing so will require an inordinate amount of work such as adding 100 spreadsheets to the list**). It will be weeded out before Step 2 begins.

## Some examples of Spreadsheets we are *NOT* looking for:

- Timesheets.
- Project Plans / Gantt charts.
- Personal tracking sheets.
- One-time documents developed for a specific project.

## Some examples of Spreadsheets we *ARE* looking for:

- List of fire hydrants.
- Performance measures.
- Potholes list.
- Street sweeping locations and times.
- PRA request history.
- Workforce centers.
- Street names.
- Libraries - locations and hours.
- List of Lists (Dataset that contains information about other datasets).
- Other cities' data portals.
- Reference data (data that provide reference to another dataset).

## Spreadsheets that are extracted from another system

If there's a report that you run from another system that ends up being put into a spreadsheet, that spreadsheet is *not a datasource*. However, please include that spreadsheet as a *dataset* in Step 2 (Sheet #2 in the Spreadsheet).

## RollUp Reports

If there's a report that you periodically prepare for another department (for example budget reports per department that roll up to a budget report compiled by FM), you do NOT need to identify that spreadsheet as a *datasource* or a *dataset*. However, we are asking you to use your judgment here -- if this is something that gets PRA'd and your department (as opposed to FM in this example) compiles the response to the PRA, or if this is something that you think may be useful to someone in another department that is not receiving the report, please put that in as a *dataset* and the system where it comes from as a *datasource*.

## Judgment Calls

There are some judgment calls required here, and if it appears your submission included too much or too little information, we will follow-up with you before reporting a status update on our inventory efforts.

## Step 2 - Dataset Brainstorm Guide.

---

There are six ways to approach dataset brainstorming:

- By Stakeholder
- By Question
- By Query
- By Osmosis
- By Example
- By Datasource type

### By Stakeholder

Review the list of stakeholders:

- Other departments
- Other divisions in your department
- Other municipalities
- County
- Council Offices
- Mayor's Office
- DCOOs / COOs
- State Gov
- Federal Gov
- Community Organizations
- Non-Profits
- Residents
- Vendors
- Media
- Management
- PRA
- Already publishing
- Media

Data flow between your department and these entities. Or sometimes they don't but you believe they should. You should have already filled out your list of stakeholders. For every datasource, go through and ask **"What data are relevant to these stakeholders from this datasource"**.

Maybe you share the datasource with them directly, but what data do they get from the datasource when it's shared? Think "lists and spreadsheets". Try to create as many as possible; we can always filter out. **You don't have to worry about determining the exact columns or column headings per dataset just yet.**

## By Question

Review the list of questions below:

- What data populate your monthly / quarterly reports
- What departmental data is currently publicly available?
- What data does your department use for internal performance and trend analysis?
- What information is published as a KPI in the budget or performance metric elsewhere?
- If you were to build a dashboard for your department, what would the metrics be and where would the data come from?
- What data are reported to federal, state or local agencies?
- What data requests do you receive under the PRA act?
- What data do other departments ask for, or you share with other departments now?
- What kinds of Open Data are similar agencies across the country publishing?
- What are the the top 10 most often run queries against the database?
- What do I think the public wants to see?
- What do I think other departments may want to pull in as part of their workflow?
- Are there data you're already sending to other departments now on a regular or per-request basis?

You don't have to respond to each one, but these will help you come up with datasets for every datasource. Think "lists and spreadsheets". Try to create as many as possible, we can always filter out.

## By Query

Talk to your department's IT person (if you have one) and try to come up with the top 5 most run queries on the datasource. Add them in as datasets!

## By Osmosis

Review portals from [Chicago](#) and [NYC](#). See what they publish and use that for inspiration.

## By Example

Review the Full Example section.

## By Datasource Type

### PDF File / Word Document

Just move that over as a dataset, and Maksim will deal with it.

## Excel File / Spreadsheet

If the Excel file contains multiple sheets, each sheet that contains "original", non-computed data is a dataset. For example, if Sheet1 contains a list of PRAs responded to, and Sheet2 Contains a subset of that that was responded to in the last quarter, only Sheet1 is a dataset. If Sheet2 on the other hand contains a list of divisions that responded to the PRAs in Sheet1, then both Sheet2 and Sheet1 are datasets.

## Access Database

This one is tricky as Access DBs can be small or huge. If you think the entire database makes sense to expose as a dataset because the tables are just not relevant without each other, feel free to put that in as a dataset. If you have stored queries that you already run against the database, those are prime dataset examples. Don't forget to try Some of the Other Brainstorming Methods Above.

## GIS Server

You will have to use your judgment here, but this should still be pretty simple. Ideally, this would be every layer that your department owns or contributes to. Don't forget to try Some of the Other Brainstorming Methods Above.

## Website / Web Based Application

Let's use our excellent OpenDSD as an example here. When you visit [OpenDSD](#), you see several options presented to you.

The screenshot displays the 'Search by Data' section of the OpenDSD website. It features a heading 'Search by Data' followed by a paragraph: 'This data goes back to the year 2003 and is for both current and historical activity. [See related data tips. \(PDF\)](#)'. Below this are three green buttons with blue text: 'Approval Search', 'Invoice Search', and 'Code Enforcement Case Search'. Blue arrows point from the text above to each button. Below this is the 'Search by Map' section with a heading 'Search by Map' and a paragraph: 'This data is for current or recent activity. [See related map tips. \(PDF\)](#)'. Below this are three green buttons with blue text: 'Discretionary Approvals Map', 'Ministerial Approvals Map', and 'Code Enforcement Case Map'. Blue arrows point from the text above to each button.

Let's break these down into datasets: The first 3 are "Approval Search, Invoice Search, Code Enforcement Case Search". The derived datasets would be accordingly:

- List of Approvals
- List of Invoices
- List of Code Enforcement Cases.

The next 3 are maps - "Discretionary Approval Map, Ministerial Approvals Map, Code Enforcement Case Map". So the datasets would be:

- Discretionary Approvals Shapefile (OR list with coordinators)
- Ministerial Approvals Shapefile (OR list with coordinators)
- Code enforcement cases Shapefile (OR list with coordinators)

Don't forget to try Some of the Other Brainstorming Methods Above.

## SAP Module

Any reports you run from the module are good dataset candidates. They should be **high value and run regularly**, especially if you have hundreds of them. Also go through the Dataset Brainstorming Instructions above. Don't forget to try Some of the Other Brainstorming Methods Above.

## SAP Business Warehouse

This one is easy. If you have pre-built reports in WEBI -- the SAP report building tool -- each of those is a dataset. If you can think of any other relevant reports as per the Dataset Brainstorming Instructions to run in WEBI add those in as well. Don't forget to try Some of the Other Brainstorming Methods Above.

## **Step 3 - Dataset Attribute Definitions**

---

### **Datasource**

What information system or database contains the data? Or what shared server or shared drive contains the data? This is automatically synced from Step 2.

### **Dataset name**

Brief descriptive name for the dataset. This is pulled in from Step 2.

### **Department/Division**

What department collects and manages the data as an official record?

### **Assigned to Complete**

Who is assigned to complete the entries for this dataset?

### **Brief description of data**

Include a brief description of the dataset. What is the purpose? What is it used for? Include key data fields if possible.

### **Frequency of data change**

At what rate does the information in the dataset change?

### **Format**

What format are the data in? e.g:

- excel
- sql
- oracle database
- pdf
- word
- etc.

### **Existing ETLs**

Are there existing database connections or extractions?

## Existing publication

Are these data already published (made publicly available) in some form or another? In a report, on our website, or on a council docket.

## Link to existing publication

If you answered yes to the previous question, please include a link to the report or document. If you don't know what it is, it's ok, don't spend too much time looking for it.

## Priority/value

What is your sense of the relative value in publishing this data? ***This is purely subjective, but will help with the initial data prioritization.***

- **High** - Existing and ongoing requests for these data; these data address pressing information needs or pain points (within or without the City); or we have heard compelling examples of how these data could be used.
- **Medium** - These data may be useful for other departments or for people external to the City; we occasionally receive requests for this information; or we have heard some examples for how these data could be used.
- **Low** - These data have unclear value for either the public or other City departments; we have never received requests for these; or we have never heard a use case for these data.

## Data Classification

How would you classify these data? ***This is purely subjective, but will help with the initial data prioritization.***

- **Public** - these data could be publicly disseminated without any concerns.
- **Protected** - these data are protected by law or regulation and can only be shared or accessed internally and per organizational procedures; OR this information includes PII.
- **Sensitive** - in its raw form these data pose security or other concerns and could be misused to target individuals.

## Data Quality Concerns

Do you have concerns about the quality of this dataset?

## PRA Frequency

If this dataset was added because it is often requested in a PRA, what is the rough frequency that it gets asked for?

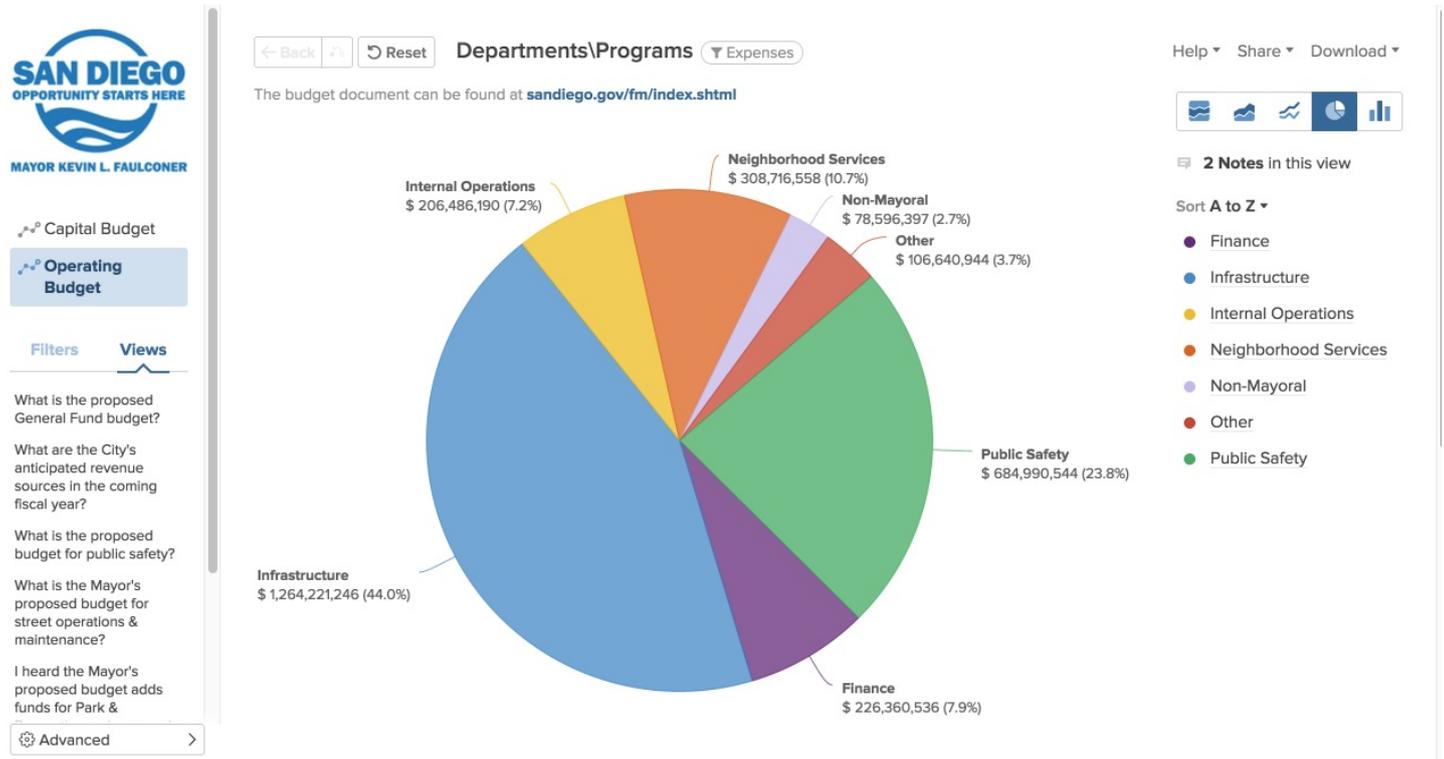
## Comments / Extra Info.

Any other information you think is relevant about this dataset.

# Other Accomplishments

## OpenGov

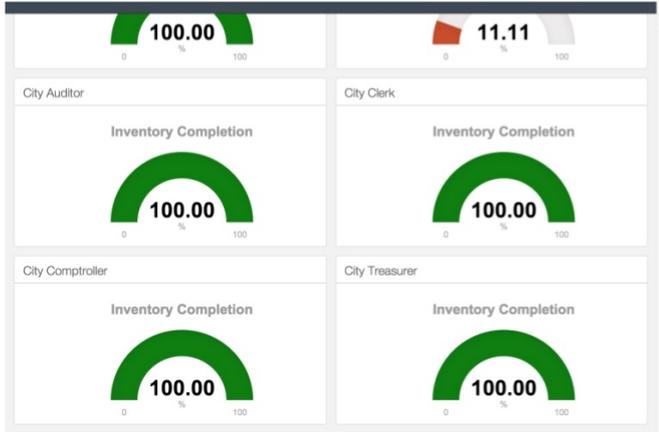
In April 2015, the City launched a public-facing budget tool, OpenGov, that to our knowledge is among the most transparent budget visualizations in the country.



Notably, our OpenGov launch took less than 4 weeks, making it the fastest large city deployment. Additionally, we are one of only a few cities that house the tool under a [subdomain of sandiego.gov](https://budget.sandiego.gov). We plan to upload FY 16 adopted as well as historical budget data in the near future. We also continue to work with the vendor to develop new functionality. In June 2015, our Director presented San Diego's story on this visualization and our Open Data program during a financial transparency webinar.

The tool is accessible at <https://budget.sandiego.gov>.

## City Clerk



Although not live yet, we repurposed the coordinator dashboard code for the City Clerk to help keep track of Records Disposition Schedules.

## City of San Diego Website Redesign



In order to ensure that it works well with the Open Data initiative, we were heavily involved in the earlier stages, and will continue to be involved in the later stages, of the City of San Diego's Website Redesign project.

## Meetings with Foreign Delegations

As the City moves along its Open Data program, we are beginning to receive global recognition for our thoughtful and insightful approaches to Open Data.

So far, we have met with delegations from the Republic of Macedonia and the Philippines to discuss Mayor Falconer's advances in Open Data and Open Government.

These visitors were brought to San Diego by the San Diego Diplomacy Council in conjunction with the International Visitor Leadership Program, the premier professional exchange program of the U.S. Department of State. Through short-term visits to the United States, current and emerging foreign leaders in a variety of fields experience this country firsthand and cultivate lasting relationships with their American counterparts.

## Meeting with the Delegation from the Republic of Macedonia

The delegation consisted of:

**Ms. Gordana DUVNJAK**

Political Journalist, MPM Media Print Macedonia, Utrinski Vesnik daily newspaper

**Ms. Ana JANEVSKA DELEVA**

Executive Director, Transparency Macedonia

**Mr. Goran KITANOVSKI**

Chief Inspector, Anti-Corruption Sector, Department for Suppressing Organized and Serious Crime, Ministry of Interior

**Ms. Neda KORUNOVSKA AVRAMOVSKA**

President, Reactor-Research in Action

**Mr. Naum PANOSKI**

Public Prosecutor, Public Prosecution Office for Fighting Organized Crime and Corruption



As citizens of a newer European country (formerly a part of Yugoslavia) the Macedonian visitors were interested in how an Open Data program and other transparency initiatives could help advance economic development and prevent corruption. We presented the benefits of Open Data and how proactively releasing public information can build trust with the community and reduce workload on City staff.

## Meeting with Philippines Delegation

The Philippines Delegation arrived with the following specific objectives:

- Enable national and local leaders to acknowledge their accountability to the office they serve as well as to the citizenry that elected them;
- Enhance civic engagement that will contribute to political, economic, and social development;
- Promote broad-based participation in the democratic process;
- Establish and promote transparency and accountability in government to eliminate the culture of corruption;
- Promote respect for the rule of law and human rights in all sectors of society.

The delegation consisted of:

**Mr. Dexter ANDO**

Special Investigator, Commission on Human Rights, Region IX, Zamboanga City

**Mr. Sherwin Prose CASTANEDA**

Assistant Director, Bureau of Import Service, Department of Trade and Industry

**Mr. Marco Dominic DE LOS REYES**

Senior Personnel Specialist, Civil Service Commission

**Ms. Norena PENONES**

Deputy Chief, Regional Intelligence Unit, Philippine National Police Intelligence Group

**Ms. Kathryn Roja RAYMUNDO**

Senior Staff Writer, Center for Media Freedom and Responsibility

**Mr. Archemides SIGUAN**

Immigration Officer, Ninoy Aquino International Airport, Bureau of Immigration, Department of Justice

**Mr. Nelson VICTORINO**

Chief of Staff, Office of Senator Grace Poe, Philippine Senate

We presented the promise and progress of the inventory process, our plans for identifying high-value data sets, and how we will release the data. The participants expressed an interest in reviewing our protocols and keeping in touch as the dialogue around freedom of information continues in the Philippines.



# Glossary

---

## About The Glossary

---

We are committed to writing all our documents in plain and readable language. There are some terms used throughout this report that may be unfamiliar or industry-specific. They are defined below, along with locations in the document where they are mentioned.

## GIS

---

A geographic information system (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present all types of spatial or geographical data.

[9.1.2. Step 2 - Dataset Brainstorm Guide](#)   [7. Community Involvement](#)   [2. Inventory Process](#)

## Git

---

Git is a free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

[0. About This Report](#)   [7. Community Involvement](#)

## GitBook

---

The toolchain (GitBook) is a tool for building beautiful books using Git and Markdown. It can generate your book in many formats: PDF, ePub, mobi or as a website.

[0. About This Report](#)   [2. Inventory Process](#)

## Hackathon

---

An event, typically lasting several days, in which a large number of people meet to engage in collaborative computer programming, usually as a competition.

[7. Community Involvement](#)

## Metadata

---

A set of data that describes and gives information about other data.

- [1. Executive Summary](#)
- [6. Prioritization](#)

## National Day of Civic Hacking

---

The goal of National Day of Civic Hacking is for residents, community groups, and government to collaborate to make their communities stronger. [National Day Of Civic Hacking](http://hackforchange.org/about/)

- [7. Community Involvement](#)
- [1. Executive Summary](#)

## Open Format

---

Data in an open format means in a convenient and modifiable form, such that there are no unnecessary technological obstacles to its use. Specifically, data should be machine-readable, available in bulk, and provided in a format with a freely available published specification which places no restrictions, monetary or otherwise, upon its use ([http://opendefinition.org/ofd/](http://opendefinition.org/ofd/)).

- [1. Executive Summary](#)
- [3. Inventory Findings](#)
- [6. Prioritization](#)

## Open San Diego

---

Group of San Diego civic-stakeholders who meet regularly to share knowledge and collaborate on projects to make our region a better place to live, work and play. [OpenSanDiego.org](http://opensandiego.org/)

- [7. Community Involvement](#)
- [5. Next Steps](#)

## Open Source

---

Open source software is software that can be freely used, changed, and shared (in modified or unmodified form) by anyone. Open source software is made by many people, and distributed under licenses that comply with the [Open Source Definition](http://opensource.org/definition).

- [0. About This Report](#)
- [2. Inventory Process](#)

## Portal

---

A web application providing access, information and sometimes visualization of data.

[0. About This Report](#)   [1. Executive Summary](#)   [3. Inventory Findings](#)   [5. Next Steps](#)

## System of Record

---

The software system that is the authoritative datasource for a given data element or piece of information.