



CITY OF SAN DIEGO

# ADOPTED BUDGET



# FY: 2016

Open Data



# Open Data

The City of San Diego is committed to the principles of open, accessible, efficient, and transparent government, as well 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 directly with each of its strategic goals and allows the City to monitor its 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>

## About the Open Data Program

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 Performance & Analytics Department 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.

# Open Data

## **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 the City describes and releases data in a way that meets Open Format specifications.

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

Understanding what data the City has is crucial to both Open Data and data management. P&A worked with departments to create an initial inventory, and now P&A needs to provide processes for how the City can stay current.

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

The City has a talented 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 six months, the Open Data Program has made significant progress on all five of its goals:

- Consistent with the policy timeline, the inventory guidelines were issued on March 9, 2015, (days before the March 31, 2015 deadline).
- Creating the City's first ever data inventory has resulted in the ability 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).
- Engagement with the civic hacking community and other interested stakeholders has been enhanced. For example, the City participated in the National Day of Civic Hacking by co-hosting three events on June 5th and 6th, 2015 (Goal 5).

## **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.

## Next Steps

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

Since the inventory is manually generated, a considerable cleaning of the data is required in order to correct duplicative and inconsistent data entry. In addition, there is a need to merge information from additional sources to validate, and possibly supplement, today's inventory. Finally, there are risks that the City's data may expose PII or cybersecurity vulnerabilities. After the inventory data has been verified and steps to mitigate these potential vulnerabilities have been taken, the data will be posted 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 it in a purely machine readable format, P&A will procure a software solution for making user-friendly visualizations of such information in Fiscal Year 2016. In addition, this software will host datasets that will be made available to the public.

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

The initial inventory largely relied on manual input, but as the City continues along the path of treating its data as valued assets, P&A 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**

P&A will continue its community outreach and engagement efforts with Open San Diego. P&A will explore new avenues for contextualizing City data, so that they can be consumed by all interested parties. P&A will pursue partnerships with nonprofits or educational institutions to analyze data for the betterment of communities and City operations.

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

P&A designed technical guidelines to operate as a living document, and it will continue to adjust as it moves 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.

# Open Data

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

After the inventory data are cleaned and analyzed, P&A will prioritize data based on the plan outlined in the technical guidelines for release. Furthermore, as Open Data is leveraged 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, P&A will be reviewing the Open Data Policy before the end of the calendar year.

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

P&A is still on schedule to begin data releases by July 1, 2016, and is looking forward to bringing in currently released data under the governance.