Open Data Report:
City of San Diego

March 14, 2014

Conducted and prepared by:

Joe Butterwick, Huron Consulting Group
Mike Brink, Huron Consulting Group
Stephen Goldsmith, Huron Consulting Group

Huron
CONSULTING GROUP

Experience. **Redefined.**™

Table of contents

I. Introduction

The City of San Diego has embarked upon an effort to develop an open data policy and approach that promotes transparency, constituent interaction, and improved municipal services. In September of 2013, Councilman Mark Kersey submitted to City of San Diego Chair of the Rules and Economic Development Committee Sherri Lightner a draft open data policy that was the product of nine months' work with interested community volunteers. On December 11, 2013, the San Diego City Council's Rules and Economic Development Committee formally requested the development of an open data policy for the City. In response, on January 10, 2014, Interim Mayor Todd Gloria created an Internal Open Data Advisory Group (Open Data Group), chaired by Deputy Chief Operating Officer for Internal Operations Jeff Sturak, to solicit input from community members and outside subject matter experts to inform the development of the open data policy. The Open Data Group is due to provide a report back to the Council by June 30, 2014.

On January 21, 2014, Huron Managing Director Steve Goldsmith and Alisha Green of the Sunlight Foundation presented an overview of open data topics to the Open Data Group and interested members of the community. This document captures a number of the points made in that presentation and the following discussion, as well as additional resources and research.

II. Creating a Successful Open Data Policy and Program

The Sunlight Foundation counts 25 state and local government policies that have been implemented since 2006. Several cities have made great progress in the implementation and advancement of open data models. These leaders include Louisville, South Bend, Philadelphia, New York, and Chicago.

A number of such policies have positive elements, but no "perfect" open data policy and program yet exists. In the experience of the Sunlight Foundation, the problem is that even well-considered legislation does not necessarily lead to the implementation of sustainable processes that lead to regular updates of datasets; therefore, the data gets "stale" and soon is of limited use. The following sections provide guidance on the development of an effective open data program that meets the civic engagement, usability, and transparency goals of such efforts.

2.1 What Data Should be Public?

A few cardinal rules should be followed when determining what data should be classified as public under an open data policy. These include:

- Set the default to open to set the expectation that all data will be subject to the policy. This creates a robust program as well as minimizing future disputes about which datasets should be subject to the program.
- Build on existing legal precedent such as public records laws to give the open data policy a firm legal standing.
- Address all government information to ensure completeness, consistency, and predictability.
- Appropriately safeguard sensitive information to protect personal information when necessary and use a balance test to assess risk when determining whether a dataset should be made public.

2.2 Making Data Public

Setting technical standards is a key element of the effort to ensure that datasets are actually useful to those who seek to access and use the data. Thus, open formats should be used that allow for bulk

download of the information of both modern and historical data. The Sunlight Foundation's guidance on both access[1] and reuse[2] is helpful. Some of the major points from that guidance follow.

The City should publish guidance as to how the data will be shared and accessed. This includes creating a data portal with clear instructions. It also means that the City should make efforts to remove barriers to access and reuse by removing arbitrary technical restrictions such as registration requirements, access fees, and usage limitations, among others. The aim should be to provide broad, non-discriminatory, and free access to data so that any person can access information at any time without having to identify him/herself or provide justification for doing so.

If information is to be truly public, there should be no license-related barrier to the public's interaction with or use of that information. Outside of data legally exempted from public use or access because of privacy or security restrictions, to be completely "open," public government information should be released completely into the public domain and clearly labeled as such. At a minimum, licenses that grant the right to use, download, and reproduce government data can be applied.

Finally, the City should take a broad view of the data that is to be made available. While it makes sense to prioritize the release of current and popular datasets, the City should consider the value of digitizing archival materials such as historical records that would add context to City information.[3] It's impossible to predict what datasets can be put to productive use by citizens, non-profits, and entrepreneurs. As was done in Philadelphia, soliciting input and/or "votes" from the public on the datasets that should be prioritized may be a helpful approach.

2.3 Developing Policy

It's important to establish proper guidance and oversight through the creation of an effective open data policy. The process of developing and adopting such a policy is an important means of securing public input and support for the program. It's important that San Diego's policy not be simply a "cut and paste" exercise from another jurisdiction.

The legislation adopted by Council should be a framework or foundation for a good open data policy, not simply a means of mandating the creation of a data portal. It should not be an attempt to anticipate all possible future developments of the program. It should be a mandate, but not an overly-prescriptive instrument that constrains the future development of the program. Generally, the ordinance should be viewed as the "floor" or "foundation" of the program, not the "ceiling".

Some of the elements that the policy should aim for include the following:

- Describing the program in clear language;
- Wielding and building on precedent such as public records laws;
- Establishing guidelines for creating and updating data inventories;
- Setting timelines for key project milestones in order to generate and sustain momentum;
- Mandating future reviews of the policy and program at periodic intervals; and
- Referencing any necessary records disposition requirements to ensure that the program is sensitive to those.

---

[1] http://sunlightfoundation.com/opendataguidelines/#open-access

[2] http://sunlightfoundation.com/opendataguidelines/#open-license

[3] http://sunlightfoundation.com/opendataguidelines/#archival-material

2.4 Selecting the Chief Data Officer (CDO)

Many cities with open data policies have found filling the position of Chief Data Officer (CDO) to be an important step in legitimizing the open data effort. There is no "one size fits all" approach to filling the position of CDO. Examples of effective CDOs can come from a more internally/service-focused background or, alternatively, a more externally/community-focused approach. If it is the right person, either model can work. However, in either case, it is essential that the individual has a direct connection to the top – an executive champion who is committed and has the organizational strength to push the open data initiative when it meets internal resistance, as it most certainly will.

The individual selected should have a familiarity and comfort level with the technology involved in open data programs, but the position is not, primarily, a technical position. It is important that the individual have solid people skills, commitment to the initiative, and an ability to articulate a compelling vision of the benefits of the program.

2.5 Building the Program

Further development of the program should be entrusted to the CDO, who should be tasked with further specifying and developing its details. Such detail may best be captured in the form of a regularly-updated guidebook that is developed under the oversight of the CDO. The guidebook developed by the City of Philadelphia is one that has value as a model.[4]

Additionally, the CDO should:
- Require that departments review their datasets and report to an oversight authority (likely the CDO) concerning the application of the open data policy to those datasets;
- Create a comprehensive data inventory and a process for routinely updating such;
- Develop a plan for, and roles and responsibilities around, updating and maintaining datasets;
- Require status reports from those individuals with responsibilities; and
- Establish a schedule for on-going progress reviews.

The CDO should also be the catalyst for ensuring that open data is leveraged by the different organizations within City government, as well as outside of it. Open data policies should allow data to be moved easily across departments with the goal of better coordinating service and leveraging citizen assistance and input in improving City services (sometimes termed "co–production" of services).

2.6 Managing the Program

Much of the success and relevance of the program will come down to the day-to-day management of the effort. A few key elements to keep in mind include:

- Prioritization is important. Decisions concerning what should be rolled out first can generate momentum for the program or, conversely, stall it. The assessment should be based upon considerations of both the utility of the data and the ease of releasing it.
- There must be a plan for regularly updating the datasets released. Stale data is the great enemy of effective open data programs. Program irrelevancy can quickly follow the public's perception that the program features stale data.

4

https://docs.google.com/document/d/1Kd4AOoRG8q18PVZ0JMusgKWJmgjrWvv3iTdKUjLEdT4/edit?pli=1https://docs.google.com/document/d/1Kd4AOoRG8q18PVZ0JMusgKWJmgjrWvv3iTdKUjLEdT4/edit?pli=1

- While in theory an open data program should represent no more of a potential liability for a jurisdiction than compliance with public records laws, the issue of City liability resulting from the use (or misuse) of data accessed through an open data program should remain a consideration and factor in the process of evaluating a dataset for release.
- The team of departmental open data coordinators should have means to assess frequently requested datasets, potentially identified by analyzing public records requests, and use that information to prioritize datasets for release. The prioritization process should also include public input.
- In order to expedite the posting of useful datasets, it's important to engage the people within City government who know the data best.
- The CDO must be able to identify the training needs of the team and address those directly or through other external resources.
- It's important to adequately staff the effort in terms of being able to fulfill timeline commitments that are publicly made. Nothing diminishes support for the program like an under-resourced effort that is unable to meet publicly-communicated milestones.
- Although there are no definitive data from other jurisdictions on this yet, it's possible that the City may need to create a position or positions to deal with specific questions from the public regarding the information available through the open data program. This is not necessarily a bad thing, however, as improving City services overall requires the inclusion of greater citizen input and participation in daily workflows. It's also possible that some City agencies could see fewer requests for more information because it would be proactively released online.

2.7   Making the Data Useful

Open data policies are a good thing in and of themselves because they promote transparency and a better-informed citizenry. However, public organizations and the community as a whole can further benefit from the analysis of organizational data by outside groups. Enabling this analysis involves putting in place helpful analytical and communication tools. Examples of such tools, approaches, and benefits include:

- Both Chicago and New York City have provided tools that allow for visualization geographically and quantitatively to enhance the ability of users to grasp and conceptualize issues like localized flooding and immunization access.
- Boston is a leader in enabling citizens to create, communicate, and use data via mobile applications to promote citizen involvement in identifying, and potentially resolving, service delivery issues.
- San Francisco presents civic questions to its citizenry and solicits their input. This approach can improve perceptions of government responsiveness as well as serve as a low cost means of generating creative responses to pressing issues.
- Washington, DC is a leader in enabling citizens to rate public services and identify issues with customer service. This has led to the development of effective feedback loops that drive rapid improvements in service.
- Generally, encouraging third party use of datasets by non-profit organizations can lead to applications that improve coordination of services across the governmental and non-profit sectors.

III. Steps Necessary to Implement in San Diego

3.1   Next Steps

By establishing the Open Data Group, the City has laid the groundwork for a successful policy development effort. Summarizing a number of the points above, the City must move forward with a policy and program that:

- Takes into account best practices;
- Sets clear timelines;
- Establishes oversight;
- Creates accountability;
- Ensures sufficient funding;
- Includes public perspectives; and
- Mandates future review.

3.2   Areas of Further Study

There are a number of issues that warrant further consideration by the City as it moves forward with its open data initiative. Those include:

- Understanding the nature and implications of inconsistencies across neighboring jurisdictions regarding open data policies;
- Assessing the financial and operational impacts of open data, as there are limited solid data on these questions at this time;
- Determining the nature and scope of the major IT infrastructure investments required to make open data a reality;
- Handling potential ADA-related concerns and challenges, though there is no indication that ADA issues have impeded the development of open data programs elsewhere; and
- Specific issues of liability, though these should not be different than those associated with existing access to public record laws.