The State of Open Data in Los Angeles County

PHASE I: A FRAMEWORK  |  MAY 2015

Dana Chinn
Sinduja Rangarajan
Brett Shears
Tess Thorman

OPENDATALA.ORG
Contents

About this report ................................................................. 1

Framework overview ....................................................... 3

Data ..................................................................................... 4

How much does a city have? And is it open?
Does a city have high quality open data?
Overall scores of eight cities in Los Angeles County

Leadership: ........................................................................ 11

Does a city have the leadership to build a robust and sustainable open
data culture?

The crossroads of data and leadership ................................. 13

Applying our framework to eight cities in
Los Angeles County ............................................................ 14

Next steps ........................................................................... 15

About us .............................................................................. 17
The purpose of this report is to propose a framework for tracking the status of open data initiatives throughout the greater Los Angeles region.

Cities and other government jurisdictions, large and small, have started many open data initiatives over the past two years. This flurry of activity, while admirable, raises several questions. Do governments have open data strategies and goals? How much progress have they made? Will their open data initiatives be sustained beyond the next round of elections?

Government data has long been “open” in the sense that governments are required to make public information, such as expenditures and property tax records, available on request. However, governments sometimes make it difficult for the public to actually get the information it seeks. And analyzing and otherwise using the data to understand government actions requires technical skills and subject–matter expertise that only a few, such as news organizations, possess.

Because a major part of an open data initiative is getting datasets online to make them accessible, we sometimes think governments believe open data is just a trendy technology project in which colorful web portals and hackathons take precedence over providing usable, relevant datasets.

We want governments to think critically about how they measure open data success and allocate scarce resources. We believe we should hold governments accountable for building a sustainable foundation, and for evaluating progress based on the breadth and quality of the datasets themselves.

For more about what open data is, go to our website at opendatala.org
OVERVIEW
Our framework is built with measurable indicators of a city’s level of open data expertise and leadership.

DATA
We blended two sets of scoring criteria to give each city an overall data score: the U.S. City Open Data Census ratings, and the detailed financial transparency criteria developed by CALPIRG, a public interest research group. Financial data is just one of many types of open data, and arguably one of the most important. Future phases of our research will include examining or developing similarly detailed criteria for other categories.

LEADERSHIP
Next, we developed criteria for assessing whether a city had the leadership infrastructure in place to integrate open data into its culture. We identified three essential components:

- An open data executive policy or city legislative action
- Full-time open data staff
- Dedicated open data funding

APPLYING THE FRAMEWORK
We rated eight incorporated cities in Los Angeles County, selecting those that already had U.S. City Open Data Census scores. We then analyzed each city’s leadership score vs. its overall data score. Our framework also incorporates insights from interviews and a survey of officials in 50 of the 88 incorporated cities in Los Angeles County.
Data: How much does a city have? And is it open?

Cities have hundreds of datasets. Each crime report filed, check written and building permit granted is kept and is public data. Robust open data portals have a wide variety of datasets that are easily available online.

The U.S. City Open Data Census (us-city.census.okfn.org), a partnership between the Open Knowledge Foundation, the Sunlight Foundation, and Code for America, is a crowd-sourced rating system based on datasets in 19 categories. The census scores cities based on a simple count of the datasets available and whether each dataset is truly “open” according to criteria such as whether a dataset is free and in a digital format.
The census score mainly indicates breadth. It is limited in its capacity both to define the overall success of a city’s open data initiative and to provide actionable information for city officials. For example:

- All of the 19 categories are weighted the same. The maximum score is 1,900, 100 for each category, e.g., crime data is given the same weight in the overall score as web analytics, and property and zoning-related datasets together are worth 700 points.

- Cities are scored for providing datasets that they neither collect nor maintain. For example, the census rewards the city of Los Angeles for publishing property tax data, which is the responsibility of Los Angeles County.

- The scoring is crowd-sourced. The scores may be incorrect, outdated, or inconsistently applied.

### U.S. City Open Data Census Categories (grouped by type)

See us-city.census.okfn.org/about for category descriptions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Category</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campaign Finance Contributions</td>
<td>Asset Disclosure</td>
<td>Lobbyist Activity</td>
</tr>
<tr>
<td>Budget</td>
<td>Spending</td>
<td>Procurement Contracts</td>
</tr>
<tr>
<td>Business Listings</td>
<td>Crime</td>
<td>Restaurant Inspections</td>
</tr>
<tr>
<td>Code Enforcement Violations</td>
<td>Construction Permits</td>
<td>Parcels</td>
</tr>
<tr>
<td>Property Assessment</td>
<td>Property Deeds</td>
<td>Public Buildings</td>
</tr>
<tr>
<td>Zoning (GIS)</td>
<td>Service Requests (311)</td>
<td>Transit</td>
</tr>
<tr>
<td>Web Analytics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
U.S. City Open Data Census scores for eight cities

COMPOSITE SCREENSHOT FROM US-CITY.CENSUS.OKFN.ORG (APRIL 20, 2015)

<table>
<thead>
<tr>
<th>City</th>
<th>Asset Disclosure</th>
<th>Budget</th>
<th>Business Listings</th>
<th>Campaign Finance Contributions</th>
<th>Code Enforcement Violations</th>
<th>Construction Permits</th>
<th>Crime</th>
<th>Lobbyist Activity</th>
<th>Police</th>
<th>Procurement Contracts</th>
<th>Property Assessment</th>
<th>Property Data</th>
<th>Public Buildings</th>
<th>Restaurant Inspections</th>
<th>Studio Requests (511)</th>
<th>Spending</th>
<th>Travel</th>
<th>Zoning (GIS)</th>
<th>WEB Analytics</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles, CA</td>
<td>0/88</td>
<td>0/858</td>
<td>0/858</td>
<td>0/210</td>
<td>4/210</td>
<td>0/210</td>
<td>4/210</td>
<td>9/210</td>
<td>9/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>0/210</td>
<td>1485</td>
</tr>
<tr>
<td>Santa Monica, CA</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>0/17</td>
<td>965</td>
</tr>
<tr>
<td>Culver City, CA</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>0/42</td>
<td>485</td>
</tr>
<tr>
<td>West Hollywood, CA</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>0/43</td>
<td>485</td>
</tr>
<tr>
<td>Pasadena, CA</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>0/44</td>
<td>480</td>
</tr>
<tr>
<td>Santa Clarita, CA</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>0/60</td>
<td>190</td>
</tr>
<tr>
<td>Bell, CA</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>0/61</td>
<td>165</td>
</tr>
<tr>
<td>Manhattan Beach, CA</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>0/62</td>
<td>165</td>
</tr>
</tbody>
</table>
The U.S. City Open Data Census gives an overview of the amount and “openness” of a city’s open data, but it doesn’t speak to the quality of the datasets. To use a dataset, a user needs to know the source, how the data was gathered, and what it includes or doesn’t - and why. Incomplete, inaccurate, or otherwise incomprehensible data leads to misuse and misinterpretation, and can result in poor decisions by government and the stakeholders who use it.

Each category of city data will have different definitions of “quality” and thus needs to be rated differently. Developing 19 sets of detailed criteria was beyond the scope for our first phase of our research, so we assessed the eight sample cities on only their “public purses,” using criteria outlined by CALPIRG in its January 2013 report on “Transparency in City Spending” (calpirg.org/reports/caf/transparency-city-spending).

CALPIRG, a non-profit public interest research group, graded 30 major U.S. cities on the availability of the budget, procurement contracts, service requests (311) and spending databases, and on the accessibility of the financial data on each city’s website. The CALPIRG criteria also gave points for the availability of past data, an essential component to understanding trends in spending. CALPIRG gave the city of Los Angeles a score of 68, or C-*. After the Open Data LA team updated the 2013 CALPIRG score, the city of Los Angeles scored 77 or a grade of C+. Culver City and West Hollywood scored slightly above Los Angeles, with 80 and 79, respectively. However, the other five cities in our sample scored below 50, indicating gaps in the availability or accessibility of their budget and spending datasets.

<table>
<thead>
<tr>
<th>CITY</th>
<th>FINANCIAL</th>
<th>SERVICE</th>
<th>WEBSITE</th>
<th>TOTAL</th>
<th>GRADE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Score</td>
<td>85</td>
<td>8</td>
<td>7</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>1 Culver City</td>
<td>73</td>
<td>3</td>
<td>4</td>
<td>80</td>
<td>B-</td>
</tr>
<tr>
<td>2 West Hollywood</td>
<td>70</td>
<td>5</td>
<td>4</td>
<td>79</td>
<td>C+</td>
</tr>
<tr>
<td>3 Los Angeles</td>
<td>72</td>
<td>1</td>
<td>4</td>
<td>77</td>
<td>C+</td>
</tr>
<tr>
<td>4 Bell</td>
<td>53</td>
<td>1</td>
<td>3</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>5 Pasadena</td>
<td>46</td>
<td>7</td>
<td>3</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>6 Santa Clarita</td>
<td>41</td>
<td>4</td>
<td>3</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>7 Santa Monica</td>
<td>41</td>
<td>4</td>
<td>3</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>8 Manhattan Beach</td>
<td>41</td>
<td>3</td>
<td>3</td>
<td>47</td>
<td></td>
</tr>
</tbody>
</table>

* Note: At USC, a 68 is a D+!
An Analysis of the CALPIRG Financial Transparency Scoring Criteria

The CALPIRG framework includes detailed scoring criteria that are relevant indicators of the quality of the datasets; it gives city officials actionable information.

Eighty-five points are given for three different types of financial data. However, 15 points are for non-financial categories: service requests (8 points), and website branding (7 points). As we’ve seen from the U.S. City Open Data Census, service requests is a separate category and warrants its own set of comprehensive criteria. Giving it only 8 points in the CALPIRG score means that a city could ignore doing it and still get a passing CALPIRG grade. Similarly, a city should be scored on its overall open data portal, and not just whether the financial data is included and accessible.

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>FINANCIAL DATA</td>
<td>44</td>
</tr>
<tr>
<td>COMPREHENSIVE ANNUAL FINANCIAL REPORT (CAFR)</td>
<td>13</td>
</tr>
<tr>
<td>MUNICIPAL BUDGET</td>
<td>28</td>
</tr>
<tr>
<td>SERVICE REQUEST</td>
<td>8</td>
</tr>
<tr>
<td>WEBSITE</td>
<td>7</td>
</tr>
</tbody>
</table>

**FINANCIAL DATA: AN ONLINE “CHECKBOOK”**
- line-by-line amounts and descriptions of every payment a city has made
- copies of contracts
- data from previous years

Points are deducted if a user has to open up individual documents to get detailed info, e.g., how much was paid, to whom, and by which department.

**BUDGETS**
- 25 points for having the current year’s budget
- Up to three additional points for previous years

**COMPREHENSIVE ANNUAL FINANCIAL REPORT (CAFR)**
- assets and liabilities
- 10 points for having the current year
- Up to three additional points for previous years

**SERVICE REQUESTS (311)**
- List of requests for city services, e.g., trash pick-up, potholes, that shows the status and completion dates
- Continuously updated
- Downloadable

**WEBSITE**
- Includes the checkbook, tax expenditure data, budgets and CAFRs in one place
- Clearly branded as financial transparency open data
For open data initiatives to live up to their power and promise, it is far more important for cities to provide the type of detailed, contextualized datasets called for by CALPIRG’s scoring criteria (see the next page) than it is to provide a large amount of raw data. Thus, each city’s overall data score in our framework consists of its CALPIRG score of only the financial data weighted at two-thirds plus its U.S. City Open Data Census score weighted at one-third.
<table>
<thead>
<tr>
<th>CITY</th>
<th>CALPIRG</th>
<th>U.S. City Open Data Census</th>
<th>OVERALL DATA SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FINANCIAL SCORE</td>
<td>INDEXED</td>
<td>SCORE AS OF APRIL 2015</td>
</tr>
<tr>
<td>Max Score</td>
<td>85</td>
<td>100</td>
<td>1,900</td>
</tr>
<tr>
<td>1 Los Angeles</td>
<td>72</td>
<td>85</td>
<td>1,485</td>
</tr>
<tr>
<td>2 Culver City</td>
<td>73</td>
<td>86</td>
<td>485</td>
</tr>
<tr>
<td>3 West Hollywood</td>
<td>70</td>
<td>82</td>
<td>1485</td>
</tr>
<tr>
<td>4 Santa Monica</td>
<td>41</td>
<td>48</td>
<td>965</td>
</tr>
<tr>
<td>5 Pasadena</td>
<td>46</td>
<td>54</td>
<td>480</td>
</tr>
<tr>
<td>6 Bell</td>
<td>53</td>
<td>62</td>
<td>165</td>
</tr>
<tr>
<td>7 Santa Clarita</td>
<td>41</td>
<td>48</td>
<td>190</td>
</tr>
<tr>
<td>8 Manhattan Beach</td>
<td>41</td>
<td>48</td>
<td>165</td>
</tr>
</tbody>
</table>
**Leadership:** Does a city have the leadership to build a robust and sustainable open data culture?

The benefits of open data - greater efficiency in responding to information requests, empowering citizens to help solve city problems, and generating new jobs, to name just a few - are long in coming.

They also require substantial and focused efforts from visionary leadership, assisted by an infrastructure that demonstrates that open data is part of a city’s DNA. An open data initiative’s progress shouldn’t be affected by changes in city leadership.

We rated cities as “high” in the level of leadership if they had at least two of the following three factors in place:

- **A policy:** An executive policy or city legislative action. The ultimate indicator that an open data initiative would be sustained would be a change in a city’s charter (if a city were a charter city).

- **Full-time staff:** At least one full-time, permanent open data staff member (i.e., not full-time temporaries or fellows assigned to open data for one or two years) who would not be reallocated to other city priorities as they flare up.

---

**Abhi Nemani**
City of Los Angeles’ first Chief Data Officer

---

continued on the next page
- **Dedicated funding**: Funds specifically allocated to implement an open data initiative, as demonstrated by a city’s use of a portal run by Socrata, Junar, OpenGov or other fee-for-service vendor that requires an ongoing development and maintenance resources.

With all three components, the city of Los Angeles was the only city in our sample to be rated “high” in leadership. The other seven have only portals, or dedicated open data funding.
Single dimensional scoring, such as we’ve seen with the U.S. City Open Data Census and CALPIRG methodologies, is limited in helping us understand where a city currently stands and what our expectations should be when tracking its progress going forward.

By plotting the level of leadership vs. the level of data quality, we argue that we can get a fair overview of:

- the pace at which we expect a city to develop open data infrastructure, and
- the likelihood that a city’s open data initiative will be sustained.

Cities without robust portals but with formal data leadership may just be starting out, and could progress quickly. Cities in this quadrant that continue to have weak portals may have leadership problems.

Cities that have robust portals and formal open data leadership are likely to progress slowly. However, changes in city leadership could reduce open data staff time and resources.

Cities that have strong open data portals, formal mandates, full-time open data staff, and dedicated funds have the best potential to become world-class open data cities.

Cities without robust portals and formal open data leadership are likely to progress slowly.
Applying our framework to eight cities in Los Angeles County

While Los Angeles has a solid foundation, Culver City’s and West Hollywood’s data scores are perhaps the result of strong but informal leadership that will continue to propel their open data initiatives. The data scores for all of the cities - and thus their quadrant positions - may change significantly when the Open Data LA team adds criteria for other datasets and website usability.

QUADRANT 1: Los Angeles
Los Angeles is the only city in our sample to have both an executive policy and at least one full-time open data staff member. The city’s data score of 83 reflects the substantial resources that both the mayor’s and the controller’s offices have spent.
The challenge for Los Angeles will perhaps be in expanding strategically based on city issues and priorities rather than just churning out more datasets and improving website usability.

QUADRANT 2: No cities
We hypothesize that no city will be in this quadrant for long. With strong leadership, a city should be able to build a robust open data portal quickly.

QUADRANT 3: Santa Monica, Pasadena, Bell, Santa Clarita, Manhattan Beach
These cities have internal champions mostly in the finance and information technology departments. Progress is uncertain, as resources are dependent on particular individuals.

QUADRANT 4: Culver City, West Hollywood
Both cities have strong informal leadership. West Hollywood has an interdepartmental “Innovation Catalyst Group” that is mentored by the city manager.
Next steps for building an open data assessment framework

Here are the key questions that the Open Data LA team will be tackling in future phases.

- **Aside from financial transparency, what datasets are the most important?**

  The datasets in a city’s open data portal should address its priorities. In other words, a portal shouldn’t just include datasets that are the easiest to prepare, the “safest” for public distribution, or the most fun or popular for hackathons.

  We believe that the promise of open data leading to more civic participation will only be fulfilled if it’s used to bring more transparency to more people around relevant topics and problems. Having only low priority datasets in portals could lead to indifference, or the perception that a government isn’t as open as it claims.

- **What are the relevant criteria to assess the quality of each type of dataset?**

  The U.S. City Open Data Census and other organizations have identified key criteria that can be used to evaluate any dataset’s usability and accessibility. However, each dataset type needs the same level of specific, detailed and weighted criteria that CALPIRG has for financial transparency datasets.
Where do the other cities and government jurisdictions in Los Angeles County fall in our preliminary framework?

Issues and problems cross many jurisdictions, especially in the greater Los Angeles region. In addition to the 88 incorporated cities in Los Angeles County, we aim to track the status of open data initiatives in Los Angeles County government, school and water districts, and other jurisdictions. Most current assessments of open data initiatives cover only the city of Los Angeles and only well-established, best practices open data cities such as New York.

Our framework could be used to establish standards across the region so that:
- the datasets can be more easily used to analyze regional issues, and
- the internal open data champions in governments can help support each other in developing best practices and more sustainable initiatives.

What role do news organizations play in the development of open data initiatives?

A data dump is no substitute for journalism. Reporters are needed to synthesize data and provide context. While open data initiatives allow non-journalists to do this as well, we believe news organizations have a new role: to hold governments accountable for proactively releasing complete and relevant datasets. And, news organizations can increase the accessibility of government information by continuing to make their own vetted and synthesized datasets available online.

The next phase of Open Data LA will include leveraging the knowledge and expertise of journalists to further develop our assessment framework. Our goal is to establish a research center to help expand journalism’s watchdog role so open data initiatives can truly thrive, and thus have more impact on government transparency and civic participation.
About us

The purpose of Open Data LA is to foster open data initiatives in the greater Los Angeles region through multidisciplinary research and other activities. We welcome your input and ideas for collaborations at OpenDataLA.org.

Dana Chinn  
Lecturer, USC Annenberg School for Communication & Journalism  
Director, Norman Lear Center Media Impact Project

Sinduja Rangarajan  
M.A. Journalism Candidate 2015

Brett Shears  
Master of Public Policy Candidate 2015

Tess Thorman  
Master of Public Policy Candidate 2016

Other USC Annenberg and USC Price faculty, staff and student researchers included Gyasi Adams, Ev Boyle, Justin Chapman, Skye Featherstone, Andrew Frantela, Liz Krane, Peter Robertson and Andrew Schrock.

Funding was provided by the USC Annenberg Center on Communication Leadership & Policy and the USC Annenberg School of Journalism.