"Lessons: Working In 'Open Data' "
by Kimberly V. Rubens, MPP 2018
Fellow, Center on Civility & Democratic Engagement
Goldman School of Public Policy

In the Spring of 2018, I was hired by the Chief Data Officer for the City of Los Angeles, Sari Ladin-Sienne, GSPP MPP 2016, to create a civic engagement framework that makes open data systems more accessible to nontechnical audiences. After writing "Open Data & Civic Engagement: Empowering Data Novices through Civic Partnerships," I took some time to reflect on global lessons that I think should inform the work of open data offices. But before I dive into these lessons, I have included a brief primer on open data systems and the identified challenges that informed my final recommendations. For a more complete history of open data systems, analysis of current open data uses, and detailed planning framework for civic partnership opportunities, check out my final report.

A brief history of open data systems

On January 21, 2009, President Obama released a Memorandum on Transparency and Open Government.[1] This memo served as a call to action for government agencies to "disclose information rapidly in forms that the public can readily find and use."[2] This directive prompted municipalities across the country to publish government data online in readily downloadable formats. Since 2014, the Mayor's Data Team in Los Angeles has created three open data portals with more than 1,000 unique datasets.

Who uses these open data systems, and what do they do with the information?

Open data portals have now existed for almost 10 years. And while these portals are incredibly useful, a survey of current users suggest that there are three types of technically trained individuals that regularly access open data systems: civic technologists, government analysts, and nonprofit employees.[3]

While these typical users are important members of the open data community, there are several planning gaps in open data content and programming that exclude nontechnical individuals from meaningful participation. First, these potential nontechnical users do not always possess the technical skills to access or analyze the data. Second, these typical, technically trained users all intend to serve nontechnical users by designing online content, tools, and products using open data that can supposedly be used by anyone. However, it is unclear based on an analysis of open data related tools and programs that the majority of these technical users regularly interact with the nontechnical

---

[1] United States, Congress, Office of Management and Budget. "President's Memorandum on Transparency and Open Government - Interagency Collaboration" https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2009/m09-12.pdf
[2] Ibid, 3.
[3] The methodology to determine these user personas is not an exact science. I read reports, talked to experts in the space, examined open data portals myself, and relied on my anecdotal understanding of who uses open data systems. These three types of users are flattened to simplify the types of users that engage in open data systems. However, among any identified type of user, one trend remains: they all possess some sort of technical training that allows them to access open data systems.

For example: The Sunlight Foundation identifies six user personas, all of whom possess intermediate to expert data analysis skills. All of New York City's user personas include savvy, digital natives or fluent digital immigrants. In both cases explicated here, as well as others surveyed, there is no explicit user persona for a digital or data novice when designing, updating, maintaining or leveraging city open data systems. The 6 user personas identified by the Sunlight Foundation are: community activist, large nonprofit project manager, small community based organization director, connector, disseminator, city staff. Sunlight Foundation User Personas: https://sunlightfoundation.com/policy/open-cities/tactical-data-engagement/equitable-neighborhoods-in-madison-wi/. The 6 user personas for NYC's open data systems are the busy bystander, equipped explorer, local liaison, influential interpreter, meticulous mapper, community champion. New York City User Personas: https://thereboot.github.io/moti-personas/

audiences they purport to serve. By excluding nontechnical users from the design process, the resulting tool or program is often irrelevant to these potential users. Finally, even if nontechnical users wanted to access open data systems, learning how to navigate open data systems requires a fairly steep time commitment.

Given this context of open data systems, below are some take-aways I mention when asked about how open data systems can be made accessible to nontechnical users.

Lesson 1: Open data teams should advocate for digital services whenever possible

When people talk about open data, they are typically referring to an open data portal. Like L.A.'s open data portal, or open budget portal, or spatial data portal. Here you can download any number of datasets uploaded by departments across the city. However, this data is in its raw form (csv, txt, shapefile) and requires at least an intermediate understanding of Excel to meaningfully analyze the data. (As well as context based knowledge to use the data in analyses or advocacy.)

Open data teams should, when possible, advocate for and invest time into digitizing government services. Examples of digitizing government services are: paying parking tickets online, applying for a construction permit, or registering your small business.

In my research, I identified two types of nontechnical users, those who are interested in learning how to access and analyze data, and those who probably just want to be able to google the answer to a question. Digital services are really aimed to help this second type of user who just wants to be able to quickly take care of government related business.

YourSTLCourts is an excellent example of a digital service. This website allows the residents of St. Louis County to enter a citation number issued by a police officer and learn which of the 81 municipal courthouses they need to go to in order to clear the ticket. In addition, this website provides text message reminders, information about the courthouse, and resources to help you pay or plea the ticket. This website was created by bringing government and nonprofit stakeholders together, to unlock 3 ticket databases covering more than 75% of St. Louis County.

Abhi Nemani, co-founder of Code For America, wrote a piece on the need for governments to focus on digital services. It is worth a read. So is this brief introduction to the U.K's Government Digital Service, a leader in digital services.

Lesson 2: Bring nontechnical users in as planning partners

The most common current users of civic open data and related programming (think Hackathons, digital literacy training events) are people that already possess either the technical training to meaningfully engage with public datasets or have the educational background as a part of their social capital to be able to easily learn the skills.

In my research, I found that the best way to make open data systems, and their programming, accessible to nontechnical audiences is to include them in the planning process. There are some truly amazing models out there that create formal and informal mechanisms to incorporate citizen-centered design.

For example, when designing YourSTLCourts, the steering committee partnered with an undergraduate class of social work students at the University of Missouri to field test the website. They surveyed 350 court users to identify needs that informed prototypes of the final website. And then spent time asking people exiting various courthouses to test versions of the website. This real-time feedback resulted in tangible changes to the current live website.

While surveying 350 court users is impressive, sometimes a survey of that scope is not always possible. In my report, I use the IAP2's Spectrum of Public Participation as a framework to outline concrete formal and informal strategies to bring nontechnical Angelenos into the planning process for the civic partnership opportunities I recommended. These strategies could be applied by any government agency looking to bring in the perspective of nontechnical users when designing open data related programming.

Below are some figures pulled from my report that detail how nontechnical audiences can be brought in as planning partners. The theory behind this spectrum of public participation is that the higher up you move, the more engaged citizens are in their government's processes.

Figure 4: IAP2 Spectrum of Public Participation

Online Tools & the IAP2 Spectrum of Participation

Figure 8: Partnership opportunities to include nontechnical users into the online tool planning process

The chart below outlines the different participation opportunities for civically-minded residents and brand new users throughout the phases of this project.

Phase	Type	Method(s)
Phase 2	Consult	1:1 conversations with civically-minded residents
		Community meetings
		Brown bag lunches, google hangouts
	Collaborate	Bring civically-minded residents and brand new users on as part of the core project team
		Develop a citizen advisory team that is periodically consulted over the course of the online tool
Phase 3	Involve	Invite civically-minded residents and brand new users to scoping meeting(s)
Phase 4	Collaborate	Include civically-minded residents and brand new users to work plan creation meetings
	Involve	Review work plan with citizen advisory committee, if formed in Phase 2
Phase 5	Inform	Provide periodic updates via social media, email list on online tool progress
	Involve	Conduct CUT testing with potential end users of the tool
	Collaborate	Conduct user testing with citizen advisory committee
Phase 6	Inform	Executing the outreach strategy
	Collaborate	Review outreach strategy with citizen advisory group or residents on project planning team
		Enlist volunteers or hire citizen advisory group, other community leaders to execute outreach strategy
Phase 7	Inform	Send periodic follow ups with information about tool usage, press coverage, success stories
	Collaborate	Include a civically-minded resident or brand new user on tool maintenance committee

Digital Literacy Training Sessions & the IAP2 Spectrum of Participation

Figure 11: Partnership opportunities to include nontechnical users into the digital literacy curriculum development process

The following chart outlines the different participation opportunities for civically-minded residents and brand new users throughout the phases of this project.

Phase	Type	Method (s)
Phase 2	Consult	1:1 conversations with civically-minded residents or brand new users to decide or develop training topics
	Empower	Poll civically-minded residents or brand new users and allow them to choose the training topic
Phase 3	Involve	Review curriculum and incorporate feedback of committed civically-minded resident or brand new user partners
Phase 4	Involve	Allow attendees to nominate people they believe would be good trainers
	Empower	Invite past attendees of trainings to lead future trainings
Phase 5	Inform	Executing training advertisement strategy
	Collaborate	Review outreach strategy with committed civically-minded residents or brand new users
		Enlist volunteers or hire civically-minded residents / brand new users to execute the outreach strategy
Phase 6	Inform	Once feedback is incorporated, email list of attendees thanking for participation and informing how feedback was incorporated
	Involve	Incorporate feedback from participants into future training materials

Lesson 3: Any good data visualization needs context

Government agencies are publishing data dashboards, and lots of them. However, I see a lot of governments throwing up graphs, charts, and performance metrics, with little context as to why this information matters. I hypothesize that governments are trying to be as transparent as possible. The data and performance management teams might understand intimately why it is important to track certain metrics. However just publishing numbers leaves the data consumer with an incomplete picture. Context is critical, and any good data visualization needs to include some sort of comparison.

For example, if a city starts publicly tracking emergency response times, there is important contextual information that should be included. Let's say a city sets a goal for an average ambulance response time of 3 minutes. In addition to tracking this, there is additional data that could help contextualize this benchmark: How did the city decide that 3 minutes should be the goal? What was the average response time when the city first started actively tracking this? How does this city's response time stack up against comparable cities?

Answers to all of these could be incorporated in an online dashboard. However, it will take a little more planning (and probably a few more data streams) to get that helpful context out there.

Lesson 4: Nontechnical audiences may need computer literacy resources as well

When thinking about how to make open data systems accessible and relevant to nontechnical users, there are a number of questions to consider: Is the open data portal easy to navigate? Are there step by step guides we can publish that walk users through the portals? Are there quick video tutorials that can be embedded to introduce new users to the open data portal in question?

However, navigating, downloading, and manipulating a dataset from an open data portal is really the second step to bringing in any new user, especially a nontechnical user. It is going to be really important to provide connections to resources that train individuals on how to navigate a data analysis software, whether it's an online spatial tool like ESRI, or an introduction to Excel class. It is amazing to have download a city's budget data, but you also want to make sure that the user will be able to analyze that dataset.

Typically, public libraries, including Los Angeles' public library system, offer introduction to computer classes that will equip data novices with the prerequisite skills to use a computer. And if open data offices want to host or support data literacy training sessions, it is important to make sure that your attendees have before the data literacy training session on how to operate a computer and work in the statistical environment you'd like to train them on.