When I ran for Mayor, I made a promise that my Administration would prioritize open government, honesty and transparency. Early in my Administration, we strengthened the role of the Office of the Inspector General and created the Chief Integrity Officer position. But, I believed we could do more.

In April 2012, I formally committed our Administration to open government and transparency by signing an Executive Order establishing the City of Philadelphia’s first Open Data policy.

It created the Chief Data Officer position and directed City employees to publish data online in an open format.

In the time since, our Administration has built a strong foundation for Open Data in Philadelphia. We have published more than 100 datasets – including high-value data like Part One Crime. We have hired a Chief Data Officer, and we created the Open Data Working Group.

Today, Philadelphia is a national leader in open data practices.

We have more work to do. This Open Data Strategic Plan will be our guide for the next phase of Open Data in this great city. It will help us reaffirm our goal of using open data to better inform citizens, improve service delivery and increase communication and data sharing between departments. It will help us maintain a citizen-centric approach and improve the way we do business as a city.

I want to thank all of the departments, agencies and individuals who have helped build our Open Data initiative. And, I am looking forward to the results this new strategy will produce in the future and the good we can achieve.

Michael A. Nutter, Mayor of Philadelphia
EXECUTIVE SUMMARY

Open data is increasingly the metric by which a city’s transparency is measured. Since Mayor Nutter established an open data policy in 2012, over 100 datasets have been released, including Part One crime incidents, contracts for professional services, licenses and violations, property assessments, and bike thefts.

But measuring progress by the number of datasets released is limited, particularly when comparing a dataset like crime incidents to ice skating rink locations, so the open data team conducted a thorough analysis and developed a system to estimate demand fulfilled by each release, and found that 36% of datasets available are of medium to very high demand.

Lessons learned from the first phase of the open data program include a need for more formal processes around prioritization and publication, and a need to understand the diverse data challenges departments have.

The next phase of open data must see it evolve into a key part of the way Philadelphia does business as a government.

This requires changing the frame of the program in three ways:

1. Renewing the focus on the general public as the ultimate customer
2. Shifting open data to be less about technology and more about departments making open data policy decisions
3. Focusing on a broader data management strategy that empowers departments to share data internally as well as with the public.

This strategic plan applies lessons learned from the first two years of the program to a new, formal process that aims to be repeatable and encompass the “big picture” of open data.

IT’S TIME FOR OPEN DATA TO GROW - BEYOND AN INITIATIVE - TO PART OF THE WAY WE DO BUSINESS AS A GOVERNMENT. WE’LL GET THERE BY FOCUSING ON THE GENERAL PUBLIC AS OUR CUSTOMER, FACILITATING DEPARTMENT OWNERSHIP, AND IMPROVING HOW WE MANAGE AND SHARE DATA IN THE FIRST PLACE.

TIM WISNIEWSKI
CHIEF DATA OFFICER, CITY OF PHILADELPHIA

THE BIG PICTURE: PHILADELPHIA & THE ROLE OF OPEN DATA

Over the last two years, Philadelphia has proven itself a leader in the open data movement. By publishing data online, city government plays a crucial role as facilitator, champion, and advocate for innovation in the city.

Open data creates opportunities for innovation around service delivery within city government and provides citizens the information they need to be more engaged. This enables government, businesses, and citizens alike to develop new solutions to complex problems. While this strategic plan captures several success stories and use cases, we are only just beginning to see the impacts of open data.

Richard Negrin,
Deputy Mayor for Administration & Coordination and Managing Director

Adel Ebeid,
Chief Innovation Officer

Richard Negrin,
Deputy Mayor for Administration & Coordination and Managing Director

Richard Negrin,
Deputy Mayor for Administration & Coordination and Managing Director

Richard Negrin,
Deputy Mayor for Administration & Coordination and Managing Director
Governments collect and use data every day to make decisions and provide services to their constituents. Data is collected to track crime incidents, report on the annual budget, and monitor the performance of schools. Most of this data is public record and subject to Freedom of Information Act (FOIA) or Right-to-Know (RTK) requests, but not everyone knows about FOIA and RTK, where to find the form, whom to send it to, or even what data is available. Open data takes the approach of proactively publishing these records on the web for people to access and analyze. Over the past decade, cities and countries across the world have implemented open data policies and laws increasing transparency, driving business, and enabling data-driven decisions.

**INCREASES TRANSPARENCY**

Government transparency is a key component of a democracy, enabling its citizens to learn what their government is doing, how it is performing, and hold it accountable. The visibility and potential to be held accountable also encourages more responsible governance.

Open data is an international trend that indicates a shift in the understanding of transparency away from a vague concept like openness and towards the more tangible and measurable medium of data releases. The transparency of a government can now be more easily measured and compared to other governments, as recently demonstrated by the Open Knowledge Foundation’s survey of US cities’ open data programs.

**DRIVES BUSINESS**

While open data is a relatively recent trend, there is already evidence that sharing government data can spur growth and innovation in the private sector. Some examples include real estate web sites using crime data to inform home buying; contractors using property violation data to identify potential customers in need of repair; and entire industries being born around the Federal Government’s release of Census and GPS data.

**RELEASING GOVERNMENT DATA HELPS BUSINESSES FIND NEW AND BETTER WAYS TO SERVE CUSTOMERS, WHETHER IT’S HELPING TO LOCATE THE NEAREST BIKE LANE OR PARCEL OF LAND FOR DEVELOPMENT. THE CITY’S OPEN DATA PORTAL CAN LEAD TO MORE ECONOMIC ACTIVITY AND IMPROVES THE QUALITY OF DOING BUSINESS IN PHILADELPHIA.**

ROB WONDERLING
PRESIDENT AND CEO OF THE GREATER PHILADELPHIA CHAMBER OF COMMERCE

**OPEN DATA 101**

**CRIME INCIDENTS (PART ONE CRIME)**

**PROPERTY ASSESSMENTS**

**LOYBYSIST ACTIVITY**

**BUILDING ENERGY USAGE**

**BIKE NETWORK**

**PROPERTY MAINTENANCE VIOLATIONS**

**BUILDING PERMITS**

**CONTRACTS (PROFESSIONAL SERVICES)**

**TOP DATA RELEASES**

**LEADS TO OPEN GOVERNMENT**

While publishing data on its own does not make “open government,” it is a critical first step towards facilitating it. Providing easily accessible information on who government is, how it works, and what information it uses to make decisions builds public trust. Informed citizens are more likely to participate in government – whether through traditional means like voting and community meetings or through more recent channels like social media and online applications, such as Fix Philly Districts, an online contest where citizens used the legislative district boundaries dataset to propose improvements to City Council during the redistricting process.\(^1\)

**BENEFITS GOVERNMENT**

Making datasets available to the public also makes them available to other government departments, which may have had difficulty accessing them otherwise, whether through fault of technology or communication. This access can enable more informed decision making and improve coordination between departments. For instance, a department that issues permits can automatically confirm there are no outstanding violations or fees owed to other departments by the applicant before issuance.

**OPEN DATA AT A GLANCE**

**DATASETS CURRENTLY PUBLISHED IN BULK, DOWNLOAD AND / OR PUBLIC API FORMAT.**

155

**CITY AGENCIES OR QUASI-GOVERNMENTAL AGENCIES HAVE RELEASED OPEN DATA.**

29

\(^1\) FixPhillyDistricts.com

\(^2\) http://us-city.census.okfn.org
Mayor Michael A. Nutter signed Executive Order 1-12: Open Data and Government Transparency in April 2012. Since then, the open data team has worked with departments across the city to pull data from various systems and publish them on the web. This work has included groundbreaking releases like Crime Incidents (Part One Crime) from the Police Department; Permits, Licenses & Violations from the Department of Licenses & Inspections; Property Assessments from the Office of Property Assessment; and many others.

Given the volume of datasets published and the number of departments participating, it was a task in itself to find the exact number of datasets released. Further, many datasets were consistently updated but others were published years ago and had fallen out of date or weren’t available in the appropriate format.

In order to evaluate the progress of the program, the open data team conducted a comprehensive analysis of published datasets, thoroughly reviewing every portal through which city data has ever been published and collecting information about the quality of each publication. The resulting compilation has been named the Open Data Census.

Since the executive order was signed in April 2012, 111 datasets have been published. Beyond that, 44 datasets were published prior to the executive order, with a total of 155 datasets from 29 departments online today.

But the analysis also pointed out that measuring progress by counting datasets alone is limited, particularly when comparing a data release like Crime Incidents to Ice Skating Rink locations. So the team began to develop a system to measure progress by estimating the demand for each dataset and comparing it to the cost/complexity of its publication. By these preliminary (and subjective) estimates, the team found that 36% of the data releases were of medium to very high demand, as illustrated by Figure 1.

In addition, by applying this estimation methodology to a list of known unpublished datasets, the team found that (a) the majority of the known datasets with low cost/complexity have been published and that (b) the majority of the remaining known datasets (unpublished) are of high cost/complexity, as illustrated by Figure 2.

The rate of progress and several particular data releases have catapulted Philadelphia among the leaders of open data nationwide. However, public demand continues to rise for other high-value datasets that have been released in cities like Chicago, Boston, San Francisco and Baltimore, such as Expenditures, Towed Vehicles, and Employee Salaries. Using the research conducted and the experience of the previous two years, the open data team has drawn several lessons and will apply them to a new strategy to take open data in Philadelphia to the next level.
WHAT WE’VE LEARNED

NEED TO CONSOLIDATE OPEN DATA PORTALS

There’s a lot of City data out there – a lot more than expected. Most of it is listed on the City’s primary open data portal, OpenData-Philly.org, but much of it is only available on other portals like PHLAPI, PHLAIM, or PASDA. This makes it difficult to find out about available datasets for those not intimately familiar with the various publication portals. If a dataset is not listed on the City’s primary open data portal, it would be reasonable for a user to assume it was not available.

NEED FOR FORMAL PRIORITIZATION PROCESS

The volume of data that exists in City government is more than any City employee currently knows. It is critical to gain an understanding of the “big picture” of what datasets exist, as well as a thoughtful method of gauging demand for each dataset – both demand from the public as well as from other departments – in order to prioritize the work of sharing data.

NEED FOR AUTOMATION AND PUBLISHING TOOLS

Sharing data should not be a manual process because (a) it creates an additional burden on already busy IT staff at departments, and (b) it quickly becomes out of date. The open data team needs to provide tools to help departments keep data up-to-date automatically and provide it in consistent, useful formats.

NEED FOR PUBLISHING STANDARDS

Many datasets have very important legal and privacy concerns coupled with serious consequences. The criteria for what can be provided in fulfilling a Right-to-Know request is not the same as what should be published online and made available through search results (for instance, the list of all registered voters). Standards around what types of information can be published and what cannot must be developed in collaboration with the Law Department.

NEED TO DEMONSTRATE VALUE TO GOVERNMENT

It can be understandably difficult for a department to prioritize sharing data with the public when it has difficulty sharing it with other departments. The process of opening data — identifying datasets, making them centrally accessible, and providing the tools to use them — not only furthers open data but facilitates countless data sharing opportunities within government. When departments can readily refer to one another’s data, they can more easily improve and enrich their own records. Dashboards and business intelligence tools enable more data-driven decisions and leads to smarter government.

IMPORTANCE OF HOW DATA IS PRESENTED

Many datasets, if presented without proper context, can be very easily misinterpreted. For instance, displaying a heatmap of vehicle crashes at a particular intersection may suggest the intersection is remarkably dangerous – until contextualized with the total number of vehicles that pass through the intersection. With government data, this often draws controversy and redirects staff resources to clarify misinterpretations of data and put out media fires. This can understandably foster reluctance to share “raw data” on its own. By being thoughtful about how data is documented and presented, the City can help users of all levels of technical expertise understand the information behind the data.

WE’RE INTERESTED IN INNOVATIVE APPROACHES TO CURBING LITTER AND INCREASING RECYCLING PARTICIPATION. BY PUBLISHING OUR DATA, WE BENEFIT FROM THE CREATIVE IDEAS FROM THE COMMUNITY THAT WE MAY NOT HAVE THOUGHT OF OURSELVES.

DAVID PERRI

COMMISSIONER, PHILADELPHIA STREETS DEPARTMENT

SHARING CRIME INCIDENT DATA IS JUST ANOTHER MEANS OF DEVELOPING PARTNERSHIPS THAT ARE DEDICATED TO ADDRESSING CRIME AND DISORDER PROBLEMS IN THE CITY OF PHILADELPHIA. OPEN DATA FACILITATES THE CO-CREATION OF PUBLIC SAFETY.

CHARLES RAMSEY

COMMISSIONER, PHILADELPHIA POLICE DEPARTMENT

DATA IN ACTION

GroundedInPhilly.org was developed by the Gar- den Justice Legal Initiative at the Public Inter- est Law Center of Philadelphia and 596 Acres, and was designed to visualize the location of Philadelphia’s vacant parcels for individuals or community groups interested in creating or preserving a community garden, market farm, or community-managed open space. The site promotes civic engagement through open data from the Department of Records, Planning Commission, Water Department, the Department of Licenses & Inspections, Office of Primary Asses- sment and the Philadelphia Redevelopment Authority.

GO TO ➜ groundedinphilly.org

GroundedInPhilly.org was developed by the Gar- den Justice Legal Initiative at the Public Inter- est Law Center of Philadelphia and 596 Acres, and was designed to visualize the location of Philadelphia’s vacant parcels for individuals or community groups interested in creating or preserving a community garden, market farm, or community-managed open space. The site promotes civic engagement through open data from the Department of Records, Planning Commission, Water Department, the Department of Licenses & Inspections, Office of Primary Assessment and the Philadelphia Redevelopment Authority.

GO TO ➜ groundedinphilly.org

Featured on PhillyMag.com, this civic applica- tion takes open data from the Police Depart- ment and translates it onto a map where a user can then specify an area of interest and view the violent crimes that have occurred there over the last three years. Developed by David Walk, this application is now a part of the Police Department’s website.

GO TO ➜ phlcrimemapper.com
CHANGING THE FRAME OF OPEN DATA

The participation and enthusiasm of departments over the last two years not only got open data in City government off the ground but put the city on the map nationwide. With that foundation in place, the open data team aims to use the lessons learned in changing the open data program to prepare for the next several years and convert open data from a special initiative in which departments participate to the way of doing business as a government in Philadelphia.

IN ORDER TO FACILITATE THIS, THE TEAM MUST FIRST CHANGE THE FRAME OF OPEN DATA IN THREE WAYS:

FIRST, there will be a renewed focus on the general public as the ultimate customer of open data. Decisions around prioritization, justification, and presentation of data releases will center around better informing citizens and improving digital services.

SECOND, open data will be less about technology and more about the departments making open data policy decisions. The open data team members are facilitators, and not evangelists: their role is to (1) convey demand by facilitating communication between those who want data and those who have it, and (2) help departments implement their release decisions in a tech-friendly way. Increased ownership by departments will help open data last beyond anyone’s individual tenure and become a standard process of government.

THIRD, the team has recognized that enabling data sharing between departments is an equally important outcome of the process of identifying and centralizing datasets. Rather than focusing strictly on potentially public data, open data will become a component of a broader data management strategy that includes private/internal data. The work of eliminating technical barriers and establishing processes (a) makes publishing data publicly more of a policy decision and (b) helps ensure long-term continuity.

In order to change the frame of open data - focusing on the general public, facilitating department ownership, and developing a broader data management strategy for the city - the Office of Innovation & Technology has aligned (1) open data, (2) civic technology, and (3) geographic information services (GIS) to create a formal process that aims to be repeatable and encompass the full lifecycle of data. This process includes greater engagement with departments to understand the data that they capture, gauging the demand for the data from the public and other agencies, planning the release of the data according to each department’s priorities, and then providing the technical support to publish the data in creative ways that are accessible to all users.

STEP 1. MEET WITH EACH DEPARTMENT
First and foremost, it is the responsibility of the open data team to reach out to data owners and communicate the changing frame of the open data strategy by meeting face-to-face with department heads and gaining a better understanding of their data challenges. In order to fulfill the role of facilitators, team members must listen to the various challenges around data that may differ from department to department and be willing to adapt and evolve the plan from what is learned.

In addition to discovering needs, these meetings serve to establish ongoing lines of communication with each department head, ideally through a nominated data coordinator with whom the open data team can work regularly on prioritization decisions.

STEP 2. CONDUCT A DATA INVENTORY
At present, there is no single list of datasets that exist in city government. Without this, it is difficult to prioritize in a consistent manner, so data releases are typically prioritized on a case-by-case basis. If such a list did exist, a department could look at “the whole picture” and make informed prioritization decisions in batch.

The open data team will work with the data coordinator designated by each department head to build this list for the department. The list, known as a “data inventory,” will contain basic information about each dataset such as its title, description, accuracy, and sensitivity (not the data itself), which will help the department make prioritization decisions.

Such an exercise will have benefits beyond open data, and thus will include datasets that may never be made public. For instance, several departments may benefit from knowing that another department has regularly updated phone numbers of property owners.

STEP 3. REVIEW AND PRIORITIZE
With an inventory of its datasets documented, the open data team can facilitate a full review and prioritization by the department, specifically:

[continued on the next page]
datasets and classify them accordingly. Many datasets will be determined shareable as-is; others will require steps be taken to alleviate privacy concerns before sharing. Some will only be able to be shared among other City departments; still others may not be able to be shared at all. Providing an opportunity to classify the department’s datasets all at once helps to ensure responsible data sharing.

GAUGE INTERNAL DEMAND

The open data team will work with the data coordinator to determine which datasets are in high demand according to the department itself. This can include an analysis of Right-to-Know requests from the public and anecdotal reports of sharing requests from other departments, both of which typically result in extra strain on department resources to extract and share the data. Further, the ability to use advanced business intelligence tools to analyze particular datasets, made possible by identifying and centralizing data, may be of high value to the department.

GAUGE PUBLIC DEMAND

An Open Data Advisory Group has been formed consisting of members representing diverse communities of data users such as academia, business, community development, technology, and others. This group will review the inventory of unpublished datasets and provide feedback to the department on which datasets are of high demand to their communities along with anecdotes to help data owners understand why. In addition to their personal involvement in their communities, this feedback will be informed by voting from the general public on the City’s open data portal, OpenDataPhilly.

Once sensitive datasets have been classified accordingly and the open data team has collected information about internal demand and public demand, the team will present the findings to the department to aid in their prioritization decisions. With the full list of datasets and informed research on the demand of each, the department will then select their highest priorities. These priorities may include datasets that need to be shared internally as well as datasets that can be shared with the public. The key is that the open data team is not setting these priorities; it is helping the department set its priorities with informed research at hand. Using data to prioritize open data, the open data team can more accurately fulfill internal and public demands. This prioritization process will be repeated regularly as demands change and datasets are released.

STEP 4. GETTING THE PRIORITIZED DATA

Datasets prioritized by the department then become part of the open data team’s pipeline of work. For each dataset, the team will work with the department’s data coordinator to implement any cleanup necessary to get the dataset to a shareable state (e.x. redactions, aggregations, rounding the address to the hundred-block level) and set up automated workflows to extract the data, perform the cleanups, and put it into a central data store, keeping it up-to-date automatically.

STEP 5. ACCESSING THE DATA

Once the dataset has been cleaned up and pulled into a central data store in an automated process, the data will be accessible in consistent ways.

INTERNAL USERS

Department staff will be able to access an internal catalog of datasets available on the central data store - particularly, (a) their own data, (b) datasets shared internally by other departments, and (c) datasets that have been shared publicly. Each dataset will be available for download or use in common business intelligence and live dashboard tools to enhance decision making. Department staff will also be able to see the list of datasets that have not yet been made centrally available along with the appropriate point of contact to obtain the data in the interim.

TECH SAVVY USERS

Any dataset pulled into the central data store that was classified as public by the department will automatically show up in the newly enhanced open data portal, OpenDataPhilly, where users will be able to (a) download it in multiple formats, (b) visualize it with charts, maps, and tables, and (c) interact with it programmatically via an API.

The list of unpublished datasets will also be available on OpenDataPhilly, where users will be able to (a) express interest, (b) provide a justification to help the department understand the importance, and (c) be directed to the right-to-know process for the dataset in the meantime.

THE GENERAL PUBLIC

Providing datasets on a central open data portal in useful formats opens a world of possibilities, but many of these possibilities require (a) a particular level of technical skill or (b) someone else with that technical skill, like a journalist or a software developer, to make the data more accessible to the general public.

In addition to being available in raw formats on the open data portal, data will be presented on the City’s web site in a contextual and easily-digestable way. Phila.gov represents the digital front door of government; it is the only URL citizens can be expected to remember. In addition to practical functionality like paying bills and applying for permits, the City’s web site is used to find out what government does, who government is, and how it works. Users should be able to find this out through data in the same way they find it out through paragraphs, images, and videos. By embedding data in creative visualizations and applications, the average user doesn’t need to know about OpenDataPhilly and can still be reached by open data.

THE FIRST PHASE OF OPEN DATA OPENED A LOT OF DOORS

with groundbreaking releases that got the program off the ground. But with over 50 departments and thousands of datasets of varying levels of complexity, the open data program needs to scale.

This requires removing technical barriers and helping departments piece together the “whole picture” of data to make informed, citizen-centric prioritization decisions. While this approach may not yield as many quick wins, once implemented it will support higher quality releases of greater frequency, and enable the program to grow and outlast individual efforts.

Follow the progress of the plan’s implementation, add your voice to the prioritization process, and see the full analysis of the Open Data Census at www.phila.gov/data.
GET INVOLVED

DISCOVER AND VOTE

Visit the upgraded OpenDataPhilly this winter to access the datasets that are already available as well as discover unpublished datasets as they are inventoried and participate in the prioritization process by expressing your interest to the appropriate department.

GO TO → OpenDataPhilly.org

DISCUSS

Join the OpenDataPhilly discussion group to have conversations with the City’s open data team, other data providers, and other users of open data.

GO TO → phila.gov/data/discuss

FOLLOW ALONG

Check out the open data team’s research and see the pipeline of datasets that have been prioritized for release by departments at the Open Data Census.

GO TO → phila.gov/data/census

BUILD

Attend the next hackathon in Philly - no technical expertise required! Meet other data users and come up with ideas for the next app or visualization that connects the general public to their city.

GO TO → ph.ly/hackathons → codeforphilly.org